Natural strategies for light harvesting in oxygenic photosynthesis: from excess light to shade

Mascoli, V.

2021

Chapter 1

The flexibility of natural light harvesting

Photosynthesis is the natural process that sustains life on Earth by converting light energy into chemical energy. It consists of a multitude of physicochemical and biochemical reactions, which are carried out by a complex and heterogeneous apparatus. Each subsequent step involves some energy loss and, as a result, the final yield of photosynthesis (solar energy to biomass) is no more than a few percent. Increasing the efficiency of light utilization by optimizing photosynthesis is highly desirable, as it can meet the increasing demand for food and clean energy. The starting point for accomplishing this challenging task is the study of the natural processes. As a matter of fact, billion years of evolution allowed photosynthetic organisms to diversify their machinery and adapt to a variety of habitats and conditions. This means that natural light harvesting is already very flexible and potentially offers a range of solutions to our problems. A deep understanding of its molecular mechanisms will hopefully make us able to trick them to our advantage.

This thesis focuses on two important aspects that limit the efficiency of light utilization in photosynthetic organisms. The first is represented by the reduced biomass yields achieved by plants in high light. Under these conditions, where the number of absorbed photons exceeds the capability of processing them, the light-harvesting machinery switches to a photoprotective state where the excess absorbed energy is dissipated as heat to avoid photodamage. Plants have evolved a number of photoprotective strategies to survive high light but, at least under some circumstances, such mechanisms can be 'overprotective' and limit plant growth excessively. Optimization of the photoprotective response can be particularly advantageous in fluctuating light conditions and was shown to represent a promising approach for increasing crop yields (Figure 1A)[1]. One of the main knowledge gaps about photoprotection concerns the molecular mechanism responsible for energy dissipation and the way this mechanism is regulated. The first half of the thesis will address these aspects by focusing on the natural tendency of the photosynthetic antennae to switch between different states that can be functional for energy harvesting and/or dissipation.

Another important restriction for crop yields is represented by the limited photosynthetically active light spectrum. Indeed, not all solar photons carry enough energy to power photosynthetic reactions and the light conversion efficiency substantially drops in the infra-red (above 700 nm, representing the so-called 'red limit'). This is why

oxygenic photosynthesis near universally relies on chlorophyll *a*, the well-known green pigment absorbing visible light, which means that the wavelengths above 700 nm, representing over 50% of the total solar irradiance, remain largely unutilized. The spectral restrictions imposed by the use of chlorophyll *a* are particularly limiting under the shade of a dense plant canopy, where most visible light is absorbed by the upper leaves (Figure 1B). Pushing the light-harvesting capacity beyond the red limit represents therefore an appealing approach for increasing biomass yields. The most intuitive way of achieving this goal is to use other pigments absorbing at longer wavelengths (or, equivalently, lower energies) than chlorophyll *a*. Interestingly, some cyanobacterial species that populate deep shaded environments were recently found to produce a different pigment, chlorophyll *f*, which makes them able to harvest far-red photons (700-800 nm)[2]. The second half of the thesis will explore the effect of the use of chlorophyll *f* on the photosynthetic performances. This knowledge is essential if we wish to introduce this pigment into other organisms, such as plants, to extend their light-harvesting capacity.

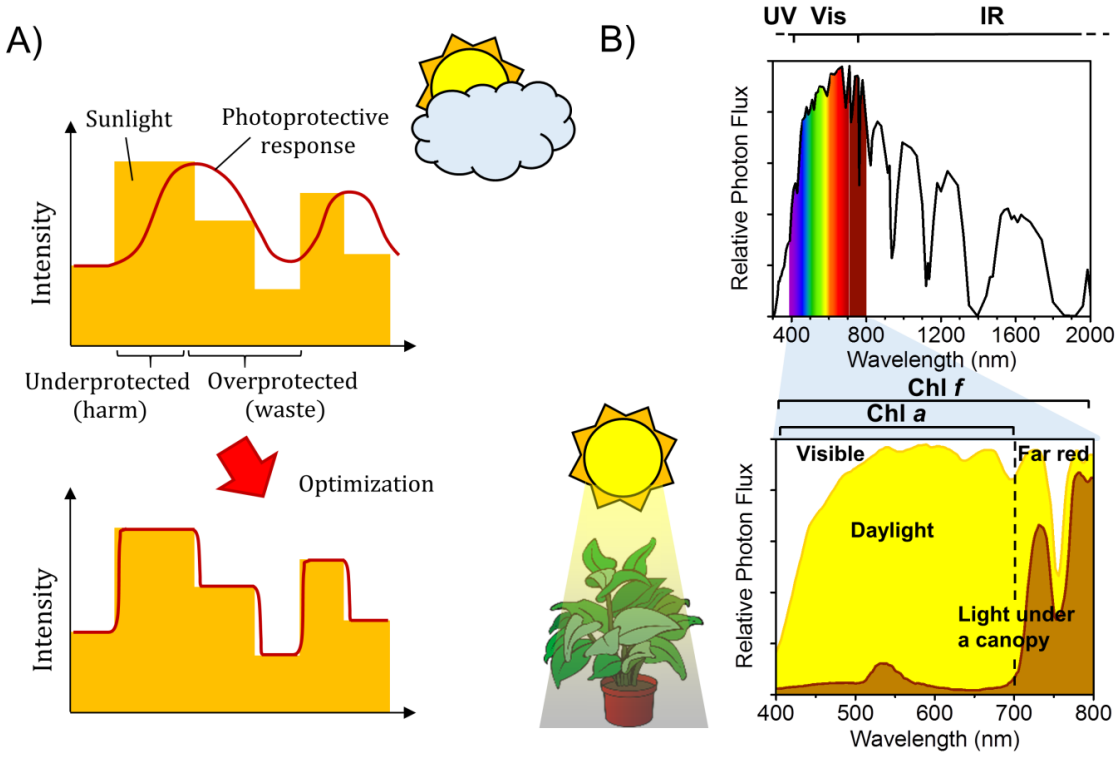


Figure 1. Light harvesting under different light conditions. A) The photoprotective response of plants (whose time course is sketched with red lines) is suboptimal under fast light fluctuations, such as on a day with intermittent clouds (different light intensities are represented by yellow blocks of different heights). The development of the response (rising red lines, top graph) might be too slow, leaving the plant under-protected when sunlight increases suddenly. Conversely, relaxation of energy dissipation (declining red lines) does not follow light intensity drops instantaneously, implying that the plants are overprotected and do not harvest light ideally for some time. A recent work has shown that over-expressing some of the photoprotective agents in tobacco leads to a prompter photoprotective response to light fluctuations (bottom graph). Consequently, light harvesting by these plants was more efficient, resulting in higher crop yields (by 15%)[1]. This figure is inspired by a work by A. Ruban[3]. B) Flux of solar photons at the Earth's surface (photons per unit surface and time, in arbitrary scale) across the electromagnetic spectrum (top) and its magnification in the visible and near IR region (bottom). The bottom graph shows the wavelength-

dependent light spectrum in full sun (daylight, yellow) and under a dense leaf canopy (brown)[4]. The latter is enriched in near IR wavelengths (above 700 nm). The spectral ranges of chlorophyll *a* and chlorophyll *f* are also indicated.

Photosynthetic light reactions

Photosynthetic organisms can collect solar photons and convert their energy into chemical energy, which they store in the form of organic compounds. The first steps of photosynthesis, the so-called light phase, are carried-out by membrane-embedded proteins, pigments and electron carriers. In the light reactions, the energy of the absorbed photons is used to produce ATP (adenosine triphosphate), the most well-known energy currency of life, and to drive electrons from water to the reducing agent NADPH (Nicotinamide Adenine Dinucleotide Phosphate). The following dark reactions, performed outside the membrane, employ ATP and NADPH to fix carbon dioxide (CO_2) into organic compounds (sucrose and starch) via a biochemical cycle named Calvin-Benson-Bassham cycle. All these reactions can be condensed in the equation:

$$6CO_2 + 6H_2O \xrightarrow{\text{light}} (CH_2O)_6 + 6O_2$$

In plants and green algae, the photosynthetic machinery is located in a specialized organelle inside their cells, the chloroplast, which is usually has a diameter of 5-10 μm[5]. Chloroplasts contain a sophisticated network of interconnected membranes known as thylakoids, where the pigments and proteins involved in the light reactions are densely packed (occupying up to 70% of the total membrane area)[6]. The thylakoid membrane separates two distinct aqueous regions, the stroma (on the outside) and the lumen (on the inside) and constitutes a vectorial system, a feature that is essential for electron and proton transfer reactions. Cyanobacteria, the only oxygenic photosynthetic prokaryotes, are the ancestors of the modern chloroplast [7] and contain analogue membrane systems.

The membrane-embedded photosynthetic machinery is organized into four major supercomplexes of proteins, pigments and other cofactors (Figure 2): Photosystem II (PSII), Cytochrome b_6f (Cyt b_6f), Photosystem I (PSI) and ATP synthase[8]. PSII, Cyt b_6f and PSI can work in series to transfer electrons from water to NADPH and protons from the stroma to the lumen. This process is globally referred as linear electron flow. Linear electron flow starts with the pigments of PSII absorbing a photon, whose energy is used to extract an electron from water and transfer it across a downstream chain of electron carriers (plastoquinone, Cyt b_6f and plastocyanin) to PSI. Here, the energy of a second absorbed photon is used to transfer the electron further to ferredoxin and ferredoxin-$NADP^+$-reductase (FNR), which ultimately reduces $NADP^+$ to NADPH. Meanwhile, protons are released into the lumen by water oxidation, and other protons are pumped from the stroma to the lumen while electron transfer steps take place. Linear electron flow requires 2 absorbed photons (one by each photosystem) for every transferred electron (from water to NADPH). Each transferred electron also results in 3 protons being pumped into the lumen. As a consequence, the complete oxidation of a water molecule (4

electrons) requires 8 photons and results in the translocation of 12 protons across the membrane. A second electron transfer pathway operates in a cycle involving PSI and Cyt b_6f only, and is therefore referred to as cyclic electron flow. While no net electron transfer is achieved, this process contributes to the build-up of the proton gradient between the lumen and the stroma. The proton gradient produced by both linear and cyclic electron flows is finally dissipated by the ATP synthase as it drives ATP production.

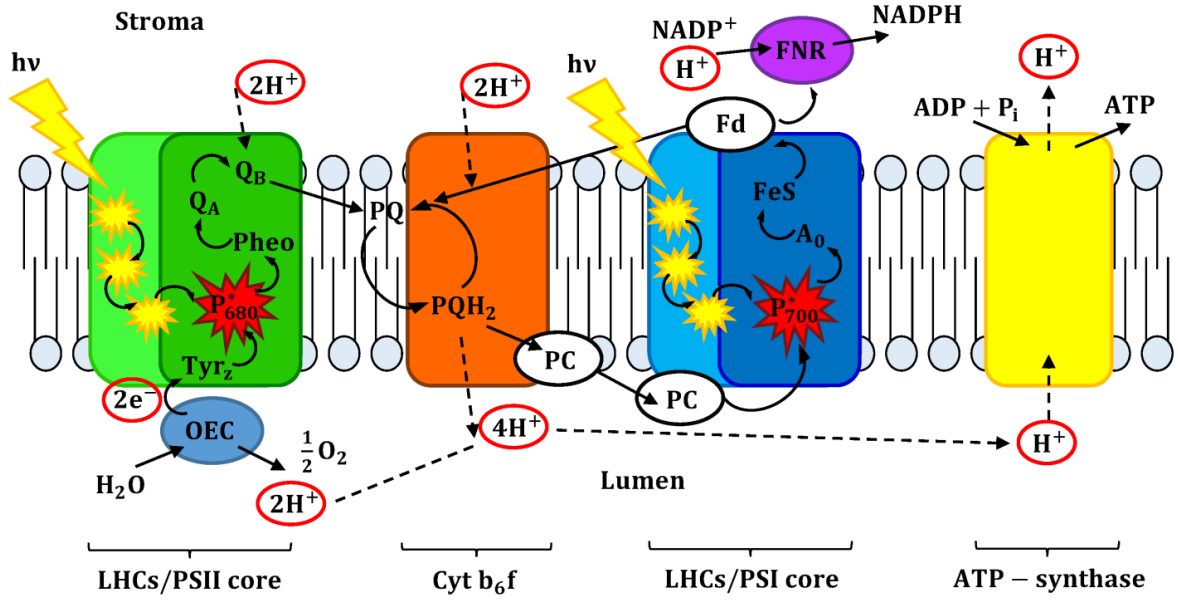


Figure 2. Photosynthetic light reactions. Scheme of the membrane-embedded protein supercomplexes involved in the light reactions. Photosystem II (PSII) and its light-harvesting complexes (LHCs) are shown in green and light green; Cytochrome (Cyt) b_6f in orange; Photosystem I (PSI) and its LHCs in blue and cyan; ATP-synthase in yellow. The electron (e^-)/proton (H^+) fluxes are indicated by solid/dashed arrows, photon (hν) absorption by yellow lightning bolts, excitations in the antenna by yellow sparks, while charge separation from the excited pigment in the reaction center (P^*_{680}/ P^*_{700}) by a red spark. Acronyms: OEC, oxygen evolving complex; Tyr, tyrosine; Pheo, pheophytin *a*; $Q_{A/B}$, immobilized/mobile plastoquinone; PQ/PQH_2, oxidized/reduced membrane plastoquinone pool; PC, plastocyanin; A_0, PSI primary acceptor; FeS, iron-sulfur clusters; Fd, ferredoxin; FNR, ferredoxin-$NADP^+$-reductase; NADP, Nicotinamide Adenine Dinucleotide Phosphate; ATP, adenosine triphosphate; ADP, adenosine diphosphate; P_i, inorganic phosphate.

This thesis focuses on the initial steps of the light reactions: 1) the harvesting of light energy by photosynthetic pigments and 2) the primary steps of its conversion into chemical energy via a process called charge separation. Light harvesting and charge separation are the two main tasks of both photosystems and are prerogative of distinct pigment pools. Charge separation, the photo-induced extraction of an electron from a chlorophyll molecule, occurs at a single central location in the photosystem by a relatively small pigment cluster called reaction center (RC). The extracted electron follows the previously described routes, whereas the oxidized pigment in the RC is re-reduced by electrons from water (in PSII) or from plastocyanin (in PSI). All the remaining pigments in the photosystems function as antennae, effectively increasing the cross-section of the RC: the absorption of a photon by a pigment in the antenna produces an excitation, which

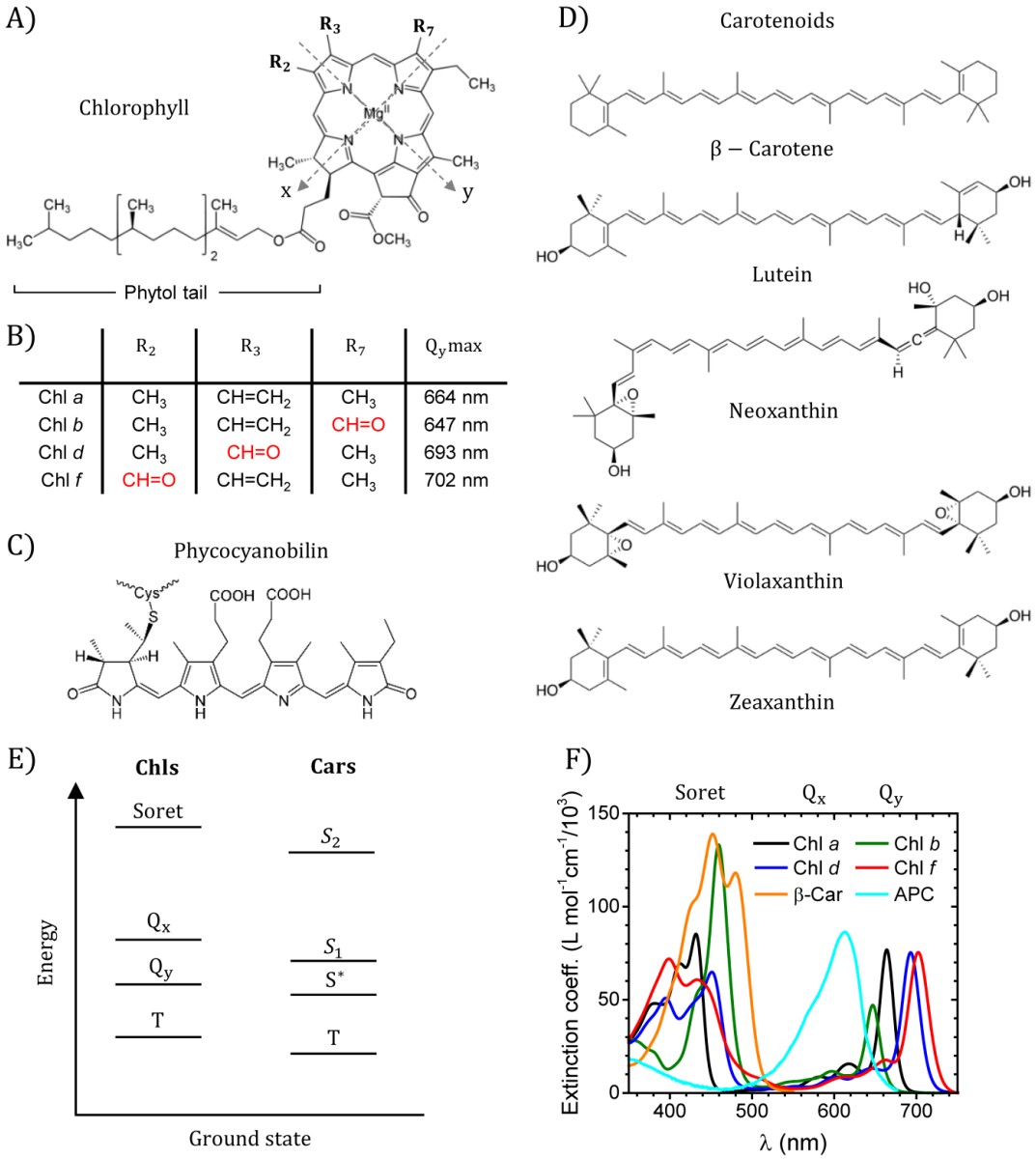
can be transferred from pigment to pigment until it is trapped by the RC, where it is used to drive charge separation. In most cases, the number of antenna pigments largely exceeds that of RC pigments, as it was experimentally observed by Emerson and Arnold in the 1930s[9] and rationalized by Duysens in the 1950s[10]. PSI and PSII supercomplexes typically consist of 1) a core complex containing the RC and a number of antenna pigments and 2) peripheral antenna complexes, whose pigments only have a light-harvesting function, i.e. they funnel excitations towards the core. While the core units of both photosystems are fundamentally similar in all organisms carrying out oxygenic photosynthesis, their antenna systems exhibit much more variability, depending on the specific environment (i.e. the spectrum and/or intensity of light available). This indicates that oxygen evolution and electron transfer reactions have relatively strict energetics that impose a specific RC/photosystem core architecture, whereas light harvesting possesses higher inherent flexibility (as long as the absorbed photons can "feed" photochemistry). The peripheral antennae of plants and green algae are membrane proteins belonging to the same gene superfamily of light-harvesting complexes (LHCs)[11]. Conversely, the major antenna complex of cyanobacteria is the water-soluble phycobilisome[12].

Photosynthetic complexes are not evenly distributed in the thylakoid membrane. In higher plants, the membranes form stacked and unstacked regions named grana and stroma/lamellae, respectively[13]. PSII is mostly located in granal thylakoids, PSI and ATP synthase in the stromal thylakoids, whereas Cyt b_6f is often assumed to be evenly distributed[14]. Despite limiting the diffusion of electron carriers from PSII to PSI, this compartmentalization, defined as lateral segregation, is probably advantageous to prevent spillover of excitations from PSII to PSI[15]. Lateral segregation appears to be absent in cyanobacteria (which also lack membrane stacking), although no consensus model has been reached on the membrane organization of these organisms[16].

Photosynthetic pigments

Pigments are organic molecules whose high number of conjugated double bonds provides them with strong absorption bands in the visible spectrum, an inescapable prerequisite for light harvesting. The main photosynthetic pigments in plants and green algae are chlorophylls and carotenoids. In addition, the phycobilisomes (PBSs) of cyanobacteria and red algae consist of phycobiliproteins (PBPs) that covalently bind pigments called bilins.

Chlorophylls and bilins. Chlorophylls (Chls) are the major pigments in oxygenic photosynthesis, as they work both in light-harvesting and charge separation. These molecules consist of a tetrapyrrole ring (1x1 nm) with a fifth isocyclic ring and a long phytol tail increasing their hydrophobicity. An Mg^{2+} ion, located at the center of the macrocycle is tetra-coordinated to its nitrogen atoms (Figure 3A). Penta-coordination of the magnesium via a protein ligand, a water or a lipid molecule is responsible for the binding of the Chls to photosynthetic proteins to form pigment-protein complexes.



	R_2	R_3	R_7	$Q_y$max
Chl *a*	CH_3	$CH=CH_2$	CH_3	664 nm
Chl *b*	CH_3	$CH=CH_2$	CH=O	647 nm
Chl *d*	CH_3	CH=O	CH_3	693 nm
Chl *f*	CH=O	$CH=CH_2$	CH_3	702 nm

Figure 3. Photosynthetic pigments. A) Structure of Chls. The crossing dotted arrows indicate the approximate directions of the Q_y and Q_x transition dipole vectors. B) Ring substituents of specific Chl molecules and maximum absorption wavelength of the Q_y transition in 80% acetone. C) Structure of phycocyanobilin. D) Structure of various Cars. E) Scheme of energy levels of relevant electronic transitions in Chls and Cars with their relative positions. The location of the Car dark states, S_1 and S^*, is shown here above and below the Chl Q_y, respectively, which is consistent with previous experimental findings on xanthophylls[18,19]. Note that other assignments have been made for other complexes/carotenoids[20,21]. F) Absorption spectra of photosynthetic pigments in 80% acetone (in units of molar extinction coefficient). APC refers to an allophycocyanin β subunit (a PBP that binds a single phycocyanobilin) from *Synechocystis* sp. (ATCC 22663). The spectrum of APC was measured in 0.01M ammonium acetate at pH 6.8[22].

The optical properties of Chls in the visible spectrum are the result of their π-electron system extending over the planar macrocycle. These molecules are green because they absorb mostly red/near-infrared (600-700 nm) and blue/violet (400-480 nm) light, while

green light is largely transmitted/reflected. The four-orbital model developed by Gouterman describes the electronic transitions giving rise to these absorption bands as linear combinations of one-electron transitions between four frontier molecular orbitals[17]: the two Highest Occupied Molecular Orbitals (HOMOs) and the two Lowest Unoccupied Molecular Orbitals (LUMOs). The resulting two low-energy (long-wavelength) excitations are referred as Q_y and Q_x (Figure 3E), as their electronic transition dipoles are oriented orthogonally in the macrocycle plane (Figure 3A). The two high-energy (short-wavelength) transitions are analogously referred as B_x and B_y or, more commonly, Soret bands.

While these features are shared by all Chl variants, other specific substituents are responsible for the differences between various Chls. Chl *a* is the most abundant and photochemically active pigment. Plants and green algae also contain Chl *b*, which is derived from Chl *a* by substitution of a methyl with a formyl (Figure 3B), and is only located in the LHCs but not in the core of the photosystems. Conversely, while most cyanobacteria only contain Chl *a,* few of them can synthesize Chl *d*[23] and Chl *f*[22,24], which possess an extra formyl group at different positions[25] (Figure 3B). In all these Chl variants, the different ring substituents influence the relative energies and electron densities of the four frontier molecular orbitals and, therefore, the energy and strength of the electronic transitions. As a result, Chl *b* has a blue-shifted (and weaker) Q_y band and red-shifted Soret bands with respect to Chl *a*, whereas Chl *d* and Chl *f* have a red-shifted Q_y transition (possibly to the near-infrared region, $\lambda > 700$ nm; Figure 3F). Thanks to these variations, different Chls absorb in different spectral regions, therefore increasing the coverage of the solar spectrum and, ultimately the light-harvesting capacity of the photosynthetic machinery.

Chls are usually bound to specific binding pockets in the protein and their spectral properties are highly sensitive to their environment, which is essential for the fine-tuning and regulation of light harvesting. First, their transition energies are sensitive to the ring geometry and to how much the Mg is displaced from the macrocycle plane[26]. Second, the Mg-ligand and other electrostatic/hydrophobic interactions with the surrounding protein residues and cofactors also influence the excitation energies[27,28]. Finally, the excited states of nearby pigments can be excitonically coupled (see the following sections)[29]. All these interactions produce spectral shifts and changes in the absorption and emission line shapes of the Chls. The study of these properties offers therefore insights in the pigment/protein structure and dynamics.

The energetic landscape of a pigment in term of its electronic states can be schematized by a Jablonski diagram[30]. The absorption of a resonant photon drives a chromophore from its ground state to an excited state resonant with that photon. Pigments' ground states are usually singlet states (S_0) and spin-conservation rules imply that photon absorption can only populate other singlet states of higher energy (S_n, n = 1,2,3,…). Excited states are highly unstable species that sooner or later relax to lower-lying states through various decay channels. Higher excited states tend to relax very quickly (usually within hundreds

of fs) to the lowest singlet excited state S_1[31]. This implies that, independent of the transition that is initially induced by photon absorption, an excited Chl will be in its Q_y (S_1) excited state after few hundreds fs[32]. This low-lying excitation can decay radiatively (i.e. by emitting a photon or, equivalently, by fluorescence), non-radiatively (heat dissipation) or by converting into a long-lived triplet (via inter-system crossing) and the average lifetime spent by the pigment in the excited state is called excited-state lifetime. In the RC of the photosystems, Chl excitations are used to drive electron transfer and therefore decay via an additional channel, photochemistry (resulting in the radical cation of the pigment electron donor). Since in a photosystem the excitations are generally formed somewhere in the peripheral antenna but can only be processed at the RC, it is immediately evident that the ideal light-harvesting pigment needs to have an intrinsically long-lived excited state. This condition is imperative for delivering the excitations formed in the antenna to the RC before they are lost via competing mechanisms. The excited-state lifetime of Chls in organic solvents is several nanoseconds[33,34], which is long enough for their excitations to cover the distance between the antenna and the RC and power charge separation with almost unitary efficiency, at least under optimal conditions. This property, together with the strong absorption bands covering a large range of the visible spectrum and the high tunability of their optical properties, make Chls the pigments of choice for natural light harvesting. The electrochemical properties of Chl *a* (and, particularly, the high oxidizing power of the Chls in the RC of PSII and the high reducing power of those in the RC of PSI) make this pigment also suitable for photochemistry[35].

Bilins are the pigments found in the peripheral antenna of cyanobacteria, the PBS[36]. Bilins have an open-chain tetrapyrrole structure (Figure 3C) that makes them more flexible than Chl macrocycles and are typically bound to the protein matrix via a thioether linkage to conserved cysteine residues. They absorb light between 500 and 670 nm (green to orange-red; Figure 3F) and are often blue-colored (therefore giving the name to cyanobacteria). Their absorption spectrum is partly complementary to that of Chls and is used to broaden the spectral range of photosynthetic active radiation (PAR).

Carotenoids. Carotenoids (Cars) are widely spread among photosynthetic organisms. They consist of a conjugated polyene chain that is often terminated by two rings (Figure 3D). Carotenoids have different lengths, ring type and isomeric forms, and can be classified as carotenes (which only contain carbon and hydrogen) and xanthophylls (which also contain oxygen). Xanthophylls are generally found in the LHCs, whereas the most important carotene, β-carotene, is mostly found in the photosystem core units[37,38]. Carotenoids have many important functions in photosynthesis: 1) they are essential for the assembly and structural stabilization of photosynthetic complexes 2) they absorb blue-green light (400-550 nm; Figure 3F) and their excitations can be promptly transferred to the Chls, implying that they are functional for light harvesting and 3) they are essential for photoprotection against oxidative damage.

The electronic structure and energetic landscape of Cars is rather distinct from those of Chls[39]. First of all, the S_0-S_1 transition can be approximated as a combination of two-

electron transitions and is, therefore, optically (one-photon) forbidden – or, equivalently, the S_1 state is optically "dark"[40,41]. The first allowed (or "bright") one-photon transition is therefore S_0-S_2, which is also responsible for the strong yellow-orange color typical of many Cars. The energy of the S_0-S_2 transition is sensitive to the conjugation length[42,43] and to the polarizability of the environment[44] (decreasing with increasing length and polarizability). Due to the flexibility of the polyene chain, this electronic transition is also strongly coupled to intramolecular vibrations/motions, resulting in rather broad absorption bands[39]. This is different from the relatively rigid macrocycle of Chls, which exhibit lower electron-vibrational couplings and, therefore, sharper absorption/emission bands[45]. Another difference with Chls, which are highly fluorescent molecules, is that Cars hardly emit any fluorescence. Upon S_2 excitation from the ground state, the excitation rapidly relaxes to S_1. This state cannot be observed via absorption from the ground state or via emission, and its energy is therefore difficult to measure and even to calculate (as its double-excitation character requires expensive and troublesome multi-reference quantum calculations). The most common way to detect Car S_1 states (and, in general, any dark state), is via its transitions to higher-lying excited states in transient-absorption measurements[46]. In this respect, Chls S_1 (Q_y) excited states are much "easier" to detect since they are fluorescent and can be directly populated via absorption. Another difference with Chl S_1, which is characteristically long-lived, is that Car S_1 states are much shorter-lived (from hundreds of fs to several ps)[39]. This property makes them excellent potential quenchers of Chl excitations in excess light (see the following sections and chapter 3). A variety of spectroscopic measurements on Cars have revealed other dark states than S_1, whose existence and origin are still debated[47]. One of them is the so-called S^* state, which has been spectroscopically distinguished from S_1 in carotenes[20,21,48,49] and many xanthophylls[18,19,50] based on the following recurring features: 1) the S^*-S_N transition is at shorter wavelengths than the S_1-S_N transition (where S_N designates a manifold of excited states accessed from S_1/S^* via direct absorption), 2) the S^*-S_N band position and line shape resemble the absorption spectrum of the carotenoid triplet and 3) the lifetime of S^* is similar to that of S_1. Some studies have assigned this band to a "hot" (vibrationally excited) ground state[51,52], even though the fact that S^* can exchange excitations with the Chls[19,21] (see chapter 3) supports the hypothesis that it represents a distinct excited state[20], possibly a different minimum on the potential energy surface of the S_1 state (i.e. it stems from the same first singlet excited state but on a different Car geometry, or conformation)[18,53]. In addition to its role in light-harvesting and quenching of Chl excitations, it has been suggested that S^* is a precursor of the Car triplet via singlet-fission processes[20].

Another important excited state of Cars is their lowest triplet state. This state is generally lower in energy than the lowest triplet state of Chls, which can be populated from the decay of Chl Q_y excitations via inter-system crossing with relatively high yields[54]. Chl triplets are very long-lived (up to 1 ms) and can react with ground state triplet oxygen ($^3O_2$) to produce singlet oxygen ($^1O_2$). The latter is extremely harmful as it can react with lipids, protein aromatic residues and purines and cause photo-damage. To protect against

this effect, Cars can accept triplets from the nearby Chls within few ns. The resulting Car triplets are much shorter-lived (few μs) and lower in energy than $^1O_2$, and are therefore unlikely to sensitize $^3O_2$. Furthermore, Cars can scavenge $^1O_2$ itself, which increases their effectiveness as antioxidants[54].

Pigment-pigment interactions: energy transfer, exciton delocalization, charge transfer

Photosynthetic complexes contain a high density of pigments that are kept in the right positions and orientation by the surrounding protein matrix in order to optimize light harvesting and its regulation. Due to this high density and connectivity, a high number and variety of pigment-pigment and pigment-protein interactions arise. Here I will give an overview of how these interactions shape the spectroscopic properties and the excited-state dynamics of the pigments, which is extremely relevant in determining the success of energy absorption, transfer and transduction in photosynthesis.

When dealing with the excited states of a multi-chromophore system (under the assumption of a single electronic transition for each pigment), two main types of interactions can be underlined: 1) the interaction between the electronic transitions of different pigments (giving rise to the so-called excitonic couplings) and 2) the interaction between the electronic and vibrational degrees of freedom within each chromophore (possibly including the surrounding protein dynamics), also called electron-phonon coupling or vibronic coupling and responsible for the absorption line-shape. For the moment we assume that the overlap in the wavefunctions of the chromophores is negligible, i.e. electron transfer between pigments is unlikely. Depending on the relative strengths of the excitonic and vibronic interactions and on the difference between the transitions energies of different pigments (also referred as site energies), two regimes can be identified[29,55]:

1) Excitonic couplings are much smaller than vibronic couplings (and/or than the energy difference between different chromophore's transitions). Under this regime, the excitations of the overall system remain localized on each pigment and pigment-pigment couplings can be treated as a perturbation to their site-specific Hamiltonian. In this view, the excitonic couplings result in the incoherent hopping of excitations between pigments. This process is referred as Förster resonance energy transfer (FRET) or excitation energy transfer (EET) and is central to photosynthetic light harvesting and to this thesis. If the two chromophores are separated enough, their transition densities are well approximated by transition dipole moments and the coupling V_{DA} between the transitions of a donor pigment D and an acceptor pigment A is reduced to:

$$V_{DA} = \frac{C_{DA}|\mu_D||\mu_A|}{R_{DA}^3}\left(\hat{\mu}_D \cdot \hat{\mu}_A - 3\left(\hat{\mu}_D \cdot \hat{R}_{DA}\right)\left(\hat{\mu}_A \cdot \hat{R}_{DA}\right)\right)$$

Where $|\mu_{D/A}|$ and $\hat{\mu}_{D/A}$ are the module and versor of the transition dipole moments of pigment *D/A* (note that the intensity of the absorption band of a pigment is proportional to the square module of its transition dipole vector), R_{DA} and $\hat{R}_{DA}$ are the module and versor of the distance vector between *D* and *A*, and C_{DA} is a screening factor which accounts for the effect of the environment. The coupling is clearly dependent on the distance and orientation between chromophores. It is therefore evident that one of the crucial roles of the protein in light harvesting is to place the pigments in the right positions and orientations and ensure energetic connectivity between them (see chapter 2 for further discussions). The energy transfer rate between pigments *D* and *A* follows from the Fermi's golden rule:

$$k_{D\to A} = \frac{2\pi}{\hbar} V_{DA}^2 U_{DA} \propto \left(\frac{1}{R_{DA}}\right)^6$$

where $\hbar$ is the reduced Planck constant and U_{DA} is the spectral overlap between the emission spectrum of the donor and the absorption spectrum of the acceptor, a term that is required for total energy conservation while the excitation is transferred. The EET rate has a very strong dependency on the inter-chromophore distance, which is why its measurement can be used to determine distances between molecules. Using some typical spectral parameters of Chls (including an intrinsic excited-state lifetime of about 2 ns, which is typical for Chls in the membrane[56–58]), it can be estimated that EET is competitive with other excited-state relaxation channels (fluorescence, internal conversion, inter-system crossing) for Chl-Chl distances shorter than 5 nm (corresponding to the so-called Förster radius). Note that an analogue expression can be derived for the reverse EET rate from A to D, by only changing the spectral overlap term (in this case between the emission spectrum of A and the absorption spectrum of D). Downhill EET from a higher energy donor to a lower energy acceptor is always more favorable than the reverse uphill process. If the EET between pigments is much faster than their excited-state decay, the pigment excited states have time to equilibrate. At equilibrium, the probability of finding an excitation on each pigment is given by the Boltzmann distribution and, the lower the temperature, the more the equilibrium favors the lowest energy states. As a result, if a multi-chromophore system is energetically equilibrated, its steady-state fluorescence spectrum becomes increasingly dominated by the longer-wavelength pigments at decreasing temperatures.

2) The opposite regime is valid when inter-chromophore couplings exceed the vibronic couplings and the pigment's transitions are relatively close in energy. In this case, it is convenient to treat the multi-chromophore system as a whole super-molecule, whose different excited states (sometimes referred as Frenkel excitonic states) are delocalized over various pigments. The properties of these excitonic states are calculated as linear combinations of those relative to the localized transitions. Under this regime, vibronic coupling can be treated as a perturbation to the electronic Hamiltonian, and is responsible for energy relaxation between different excitonic states, as described by Redfield theory[59]. A more refined

version of this approach is the so-called Modified Redfield theory[60], where only the off-diagonal component of the vibronic coupling is treated perturbatively, whereas the diagonal component is treated more accurately. This way, a more appropriate line shape for the various excitonic transitions can be recovered. These techniques are used in chapter 2 of this thesis to resolve the excitonic structure of a LHC of plants, CP29.

Obviously, the above-described regimes only represent extreme limits. Life is, as usual, much more complicated and photosynthetic complexes are often found in the intermediate regime where excitonic couplings, vibronic couplings and differences in the pigment excitation energies are in the same range. Under these conditions, the previously introduced models are not always reliable and more sophisticated theories have been introduced[61,62], which are however more expensive and therefore difficult to apply to large systems consisting of several interacting chromophores. It is also worth mentioning that the protein disorder affects the pigment energetics on different timescales, from femtoseconds to seconds or more[63]. While the fastest timescales of these dynamics contribute to the homogeneous broadening of the spectral line shapes, those significantly slower than the excited-state lifetime can be regarded as "static". This type of disorder is manifested as inhomogeneous (Gaussian) broadening of the electronic transitions and has the effect of further reducing the extent of exciton delocalization. A practical way of modeling photosynthetic complexes and account for the effect of protein disorder is to consider excitations to be possibly delocalized only over small clusters of strongly coupled pigments, whereas energy transfer between these clusters, interacting more weakly, can be reasonably described by Förster theory[64].

When two pigments approach each other more closely and their respective wave-functions display a significant spatial overlap, also electron transfer can take place between them. Electron transfer from/to an excited state is generally more favorable than from/to the corresponding ground state, as an excited pigment possesses a higher energy electron and a lower energy hole created by the electronic transition. This increased reducing power of the Chl excited states is the driving force for charge separation in the RCs of PSII and PSI. Based on Marcus theory[65], the rate of an electron transfer reaction between a donor D and an acceptor A at the temperature T is given by:

$$k_{DA} = \frac{2\pi}{\hbar} V_{DA}^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(-\frac{(\Delta G + \lambda)^2}{4\pi k_B T}\right)$$

where V_{DA} is the electronic coupling between the acceptor and donor molecules, k_B is the Boltzmann constant, ΔG is the free energy difference between the states in their equilibrium geometries after and before electron transfer, while λ is the reorganization energy, i.e. the energy of the initial electronic state at the equilibrium geometry of the final state. The coupling term depends on the spatial overlap of the wave-functions of the donor and acceptor molecules, and is therefore proportional to:

$$V_{DA} \propto e^{-R_{DA}/R_{VdW}}$$

Where R_{VdW} is the sum of the Van der Waals radii of *D* and *A*. The distance dependency for electron transfer reactions is even more severe than for EET processes, which makes electron transfer competitive only at quite short distances, typically 1 nm or lower. A similar exponential dependency on the inter-chromophore distance also applies to triplet transfer reactions. A general expression for electron transfer rates in biological systems has been determined empirically, predicting the fastest possible reaction to happen in about 100 fs[66]. Typical timescales for charge separation in photosynthetic RCs are in the order of 1 ps or more[67–69].

It is worth mentioning that, in a multi-pigment system, Frenkel excitonic states and charge transfer (CT) states are both parts of the overall energy landscape. The former are pigment-neutral states, resulting from the combination of excited states where the excited electron and hole reside on the same molecule. The latter states have the excited electron and hole localized on different pigments, implying that the net charge of some pigments is not zero. For such a reason, CT states display larger vibronic couplings, which are manifested by the increased spectral broadening and bathochromic shifts[70,71]. Electronic transitions to pure CT states are only weekly allowed, which makes them optically dark and, therefore, more difficult to observe spectroscopically (as they hardly absorb or emit any photon). In some pigment configurations, however, CT states can partially mix with excitonic states and borrow dipole strength from them, meaning that they can be detected via absorption or fluorescence[72–74]. These mixed Frenkel-CT states of Chls have been used to explain the broad red-shifted fluorescence spectra typical of the LHCs of PSI of plants and algae[73]. Due to their character of (at least partially) dark states and their possibly increased non-radiative decay rates, Chl-Chl and Chl-Car CT states have been also proposed as candidates for quenching of Chl excitations in excess light[72,74] (see following sections).

Photosystem II core: architecture and charge separation

The PSII core complexes of plants, algae and cyanobacteria are all very similar[38,75,76] and the highest resolution structure available so far (1.8 Å) is that from the cyanobacterium *Thermosynechococcus elongatus*[75] (Figures 4A and 4C). The physical center of the PSII core consists of a heterodimer of proteins D1/D2 (also named PsbA/PsbD), that together bind the 6 Chls *a*, 2 Pheophytins *a* (Pheo; Pheo *a* is Chl *a* where the central Mg is absent and two nitrogen atoms are protonated) and two plastoquinone (Q) molecules of the RC. On the lumenal side, D1/D2 also binds the water-splitting manganese cluster responsible for water oxidation and oxygen evolution.

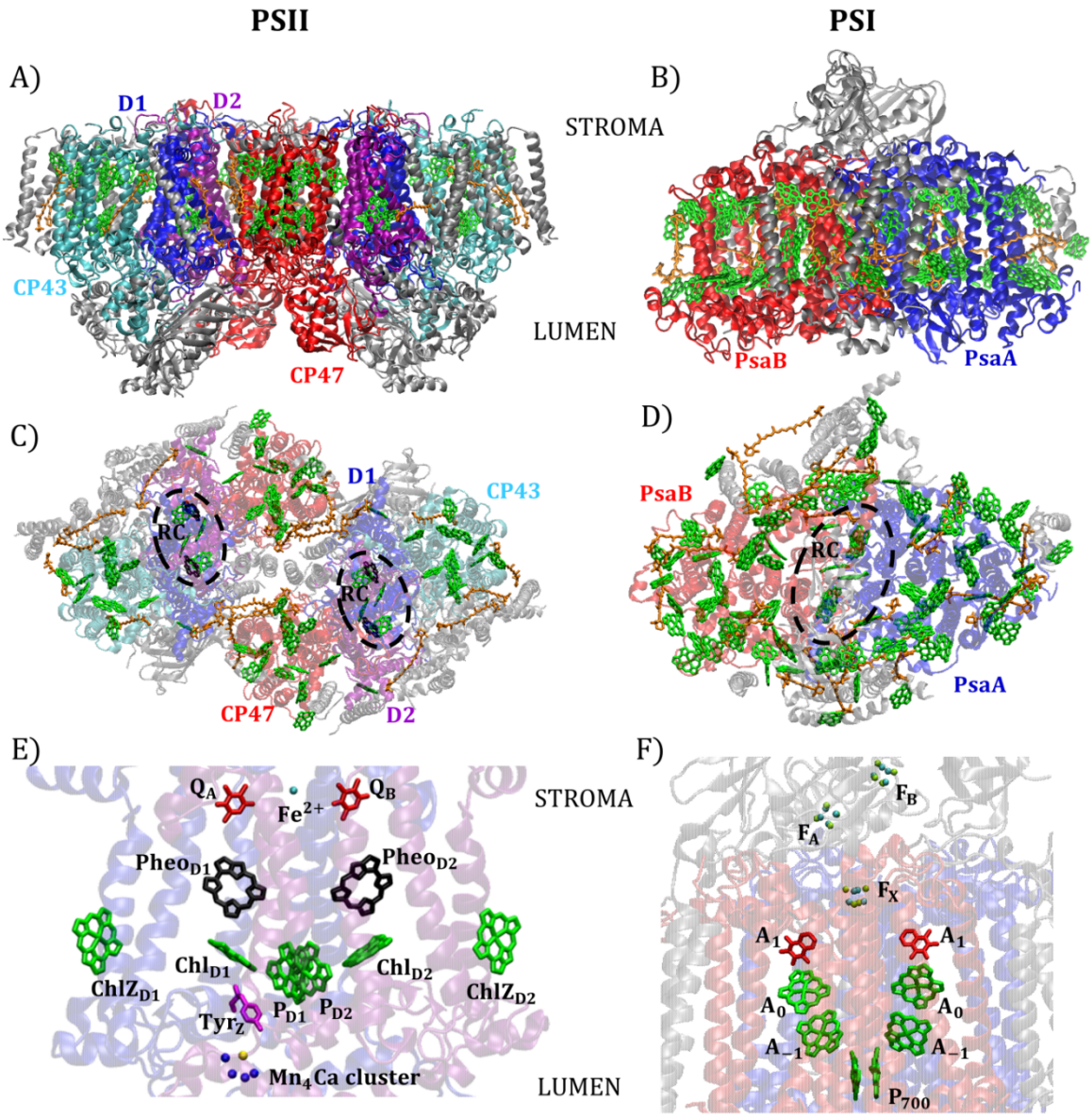


Figure 4. Structure of photosynthetic core complexes. A) Side view of the PSII dimeric core complex from *Thermosynechococcus vulcanus*[75] (Protein Data Bank (PDB) code: 3WU2). Chl *a* molecules are shown in green, β-carotene molecules in orange. Chl-binding protein subunits are shown in different colors: blue (D1), purple (D2), red (CP47), cyan (CP43). All other subunits are in grey. B) Side view of the PSI monomeric core complex from *Synechococcus elongatus*[77] (PDB code: 1JB0). PsaA and PsaB subunits are highlighted in blue and red, respectively, while all other protein subunits are in grey. C,D) View of dimeric PSII (C) and monomeric PSI from the stroma. For a better visualization of the pigment molecules, the protein is shown in transparent colors. The RC pigments are circled with a black dashed line. E,F) side view of the RC pigments of PSII (E) and PSI (F) Chl a molecules are in green, pheophytin *a* in black, quinone molecules (plastoquinone, $Q_{A/B}$, for PSII) and phylloquinone (A_1, for PSI) in red, the manganese (Mn_4Ca) cluster in blue/yellow, tyrosine Z in magenta, the iron (II) ion in cyan and the iron-sulfur (F_i) clusters in cyan/lime. For clarity, only chlorin rings of the Chl pigments are shown.

The RC pigments (except two of the Chls *a*, which are located at the periphery of D1/D2) form two pseudo-symmetric branches and electron transfer involves only one of them[78,79] (Figure 3E). Photochemistry starts from an excited state, $*P_{680}$ (named after the peak

wavelength of its absorption band), that likely involves several RC pigments, as they are similarly close both in distance and energy terms. This condition gives rise to a distribution of charge separation reactions, though the prevailing one seems to be electron transfer from Chl_{D1} to $Pheo_{D1}$, as Chl_{D1} is slightly lower in excitation energy (Chl_{D1} is therefore referred as primary electron donor)[80–82]. This first charge separation step results in the formation of a radical pair, $Chl_{D1}^{\bullet+}Pheo_{D1}^{\bullet-}$. The positive charge on Chl_{D1} is filled by an electron from the special pair of Chls P_{D1}/P_{D2}, whereas the electron further translocates from Pheo to an immobilized plastoquinone (Q_A), and then to a mobile plastoquinone (Q_B)[83]. The latter, after receiving a second electron (and two protons from the stroma) is released into the membrane as QH_2 and is replaced by a new oxidized Q. Finally, the oxidized special pair, which is a very strong oxidant, receives electrons from water via the manganese cluster and an intermediate tyrosine residue named Tyr_z[84].

D1/D2 are flanked by the core antennae CP43/CP47 (also named PsbC/PsbB), binding 13/16 Chls *a*, respectively, as well as several β-carotene molecules[38,75]. Most of these Chls absorb and emit photons in the 670-680 nm range[85,86]. Besides D1/D2/CP43/CP47, the PSII core complex contains several smaller pigment-less subunits, and it naturally occurs as a dimer. In plants and green algae, the dimeric core is surrounded by a variable number of LHCs, which will be described later in this chapter.

Electron transfer from the primary donor is believed to take place in a few ps, or even on a sub-ps timescale[69]. PSII cores, however, need a substantially longer time, on average, to process an excitation formed in the antenna by performing photochemistry. The latter event marks the end of any Chl excitation, implying that no fluorescence is emitted by PSII afterwards (this occurrence is sometimes referred as photochemical quenching). From an experimental point of view this means that the average time needed by PSII to complete charge separation corresponds to the Chl excited-state lifetime measured via time-resolved fluorescence. Based on its architecture, the fluorescence lifetime of PSII can be seen as the sum of two contributions: (i) the average time required for an excitation to reach the RC from the antenna for the first time (migration or diffusion time) and (ii) the time required by the RC to complete charge separation (trapping time). The Chls of CP43/CP47 are spatially separated from those of the RC to prevent oxidative damage (Figure 3C), which reflects into a relatively slow diffusion of the excitations from the antenna to the RC pigments (tens of ps)[87]. In addition, there is evidence that energy transfer between the antenna and the RC, as well as primary charge separation, are partially reversible (i.e. the separated charges can recombine to form the original excited states)[80,88]. This is summarized by saying that PSII RCs are shallow energy traps. As a result, excitations can move back and forth between the antenna and the RC and/or the radical pairs can recombine multiple times before charge separation is successfully completed. This is why the average fluorescence lifetime of intact PSII cores (τ_o) can rise up to 100 ps[89–91]. A common way to experimentally determine the maximal photochemical yield of PSII (Φ_{PSII}) is to compare its fluorescence lifetime when the RCs are open (i.e. the rate of charge separation is maximal) or closed (i.e. the rate of charge

separation is zero). These two extreme cases are achieved by fully oxidizing or reducing the electron acceptor side of PSII, i.e. plastoquinone. The fluorescence lifetime of Chls in the membrane in the absence of photochemistry (τ_c) is about 2 ns, implying $\Phi_{PSII} = 1 - \tau_o/\tau_c \approx 95\%$, which is pretty remarkable.

Photosystem I core: architecture and charge separation

The PSI supercomplexes of green eukaryotes and cyanobacteria have quite different organizations. The PSI supercomplex of plants and algae has a monomeric core surrounded by a number of LHCs[37,92,93] (Figure 5B), which are traditionally designated as LHCas, to differentiate them from those typical of PSII, known as LHCbs (note that, however, PSI can also bind LHCII, which consists of LHCbs). The PSI of cyanobacteria is usually found as a trimer[77,94] (or, in some species, a tetramer)[95] of core units that are, individually, rather similar to the PSI core of photosynthetic eukaryotes. One of the important differences between them is that the PSI core of plants possess two additional protein subunits that prevent the possibility of trimerization and stabilize the binding of the LHCas and LHCII[94,96].

The PSI core typically binds about 100 Chls *a* and 20 Cars (mostly β-carotenes)[37,77,93]. The vast majority of PSI core pigments, including those forming the RC, are bound to the two largest central protein subunits, PsaA and PsaB (Figures 4B and 4D). The RC, shared by these two units, is located at the center of the supercomplex and consists of 6 Chls *a*, 2 phylloquinones (or menaquinones), and 3 iron-sulfur [4Fe4S] clusters (Figure 4F). The 6 Chls and two quinones are organized in two pseudo-symmetric branches and, different from the PSII RC, both branches are nearly equally active in electron transfer[97,98]. The oxidizing species of the PSI RC, i.e. the special pair of Chls P_A/P_B, is named P_{700} after its absorption maximum wavelength. Currently, two models for charge separation have been proposed for PSI RCs. In one case the primary donor is $*A_{-1}$, transferring an electron to the nearby A_0 to form the $A_{-1}^{\bullet+}A_0^{\bullet-}$ radical pair first, followed by electron transfer from P_{700} to $A_{-1}^{\bullet+}$ to form a second radical pair, $P_{700}^{\bullet+}A_0^{\bullet-}$[98]. In the second model, $*P_{700}$ is the primary donor and transfers an electron to A_{-1} (forming $P_{700}^{\bullet+}A_{-1}^{\bullet-}$), followed by electron transfer from $A_{-1}^{\bullet-}$ to A_0 to produce the second radical pair, $P_{700}^{\bullet+}A_0^{\bullet-}$[99,100]. Both models result in the same final radical pair but identifying the primary donor (and, therefore the first intermediate radical pair) is complicated by the fact that the energetics of the A_{-1} and A_0 pigments are currently unknown and the two successive steps are extremely rapid and difficult to separate temporally. The primary charge separation steps leading to the formation of $P_{700}^{\bullet+}A_0^{\bullet-}$ take place in 1-2 ps[68]. From $A_0^{\bullet-}$, the electron is further ceded to the phylloquinone, and from there to the [4Fe4S] clusters, which finally reduce ferredoxin. $P_{700}^{\bullet+}$ is reduced by the electrons coming from $PSII/QH_2/Cyt\ b_6f$ via plastocyanin.

All other Chls surrounding the RC pigments have a mere antenna function and the majority of them emits at 680-690 nm[101,102]. Their concentration is relatively high, which is why most of these Chls form strongly coupled clusters of several pigments[77]. This is probably the reason behind the presence of few peculiar Chl spectral forms (the so called

"red forms") carrying red-shifted absorption and emission spectra in PSI cores (as well as in the LHCas). Chl red forms most likely arise from the mixing of a CT state with a Frenkel exciton state in a strongly interacting Chl dimer/multimer[73,74]. In most organisms, the presence of Chls with $\lambda > 700$ nm represents a distinct signature of PSI (as opposed to PSII) in fluorescence measurements. This does not mean, however, that some of the PSII-related components cannot give rise to a red-shifted emission under some circumstances, as will be clear in chapter 4. The exact function of PSI red forms is not entirely clear yet, though they have been proposed to expand the PAR to the far-red spectral region, to concentrate energy at strategic locations, or to help in photoprotection[101,103–105]. These low-energy Chls are favored by Boltzmann distribution and can easily gather excitations from the remaining antenna pigments. Since the red Chls are lower in energy than the RC pigments (that have $\lambda \leq 700$ nm), their excitations require some thermal energy from the environment to reach the RC via uphill EET. Typically, this is possible at physiological temperatures but not at cryogenic temperatures. Evidently, this uphill energy transfer process represents a limiting step and, therefore, the more red Chls (and the lower their energy), the longer it takes for the RC to trap excitations from them and power photochemistry[68]. The result is that the overall trapping time in PSI cores rises from about 15 ps in the absence of red forms[106], to > 50 ps in the complexes that contain the most red-shifted Chls *a*[107,108]. Despite this effect, charge separation by PSI is faster and more irreversible than in PSII. Due to the higher reducing power of the acceptor side of the electron transfer chain, PSI is in fact a very strong photochemical trap with a quantum efficiency approaching 100%, which makes it the most efficient light-driven redox enzyme in nature[109].

In line with this red-shift in the antenna pigments, the RC of PSI (absorbing at 700 nm) is also red-shifted with respect to the RC of PSII, which only absorbs at 680 nm (i.e. charge separation by PSII requires more energetic photons). These two wavelengths (680 nm for PSII and 700 nm for PSI) represent the so-called red limit of oxygenic photosynthesis. This strict energetic requirement is believed to be necessary for increasing the oxidizing power of PSII and allow water splitting. Indeed, the energy required by photochemistry in non-oxygenic photosynthetic bacteria is significantly lower (λ up to 900 nm, or even longer), but the latter type of RCs are not able to extract electrons from water[35].

Light-harvesting complexes and their organization in the photosystems

In plants and green algae, LHCs are a superfamily of nuclear-encoded membrane proteins binding Chl *a*, Chl *b*, carotenoids and a phospholipid molecule[11,110]. This gene superfamily also includes some stress-induced proteins of cyanobacteria (such as HLIPs and ELIPs, which are considered as the ancestors of the family)[111,112] and green eukaryotes (such as LHCSR and PsbS)[113,114].

The genes codifying for the LHCs of PSI and PSII are called LHCas and LHCbs, respectively. In higher plants, three monomeric LHCs, CP29, CP26 and CP24 correspond to the genes Lhcb4, Lhcb5 and Lhcb6, respectively, and are sometimes referred as minor

antennae. The other three genes Lhcb1, Lhcb2 and Lhcb3, codify for the units of the trimeric complex LHCII, which is the major LHC and also the most abundant membrane protein on earth. The four LHCas of plants, corresponding to genes Lhca1-4, are naturally organized as dimers[115].

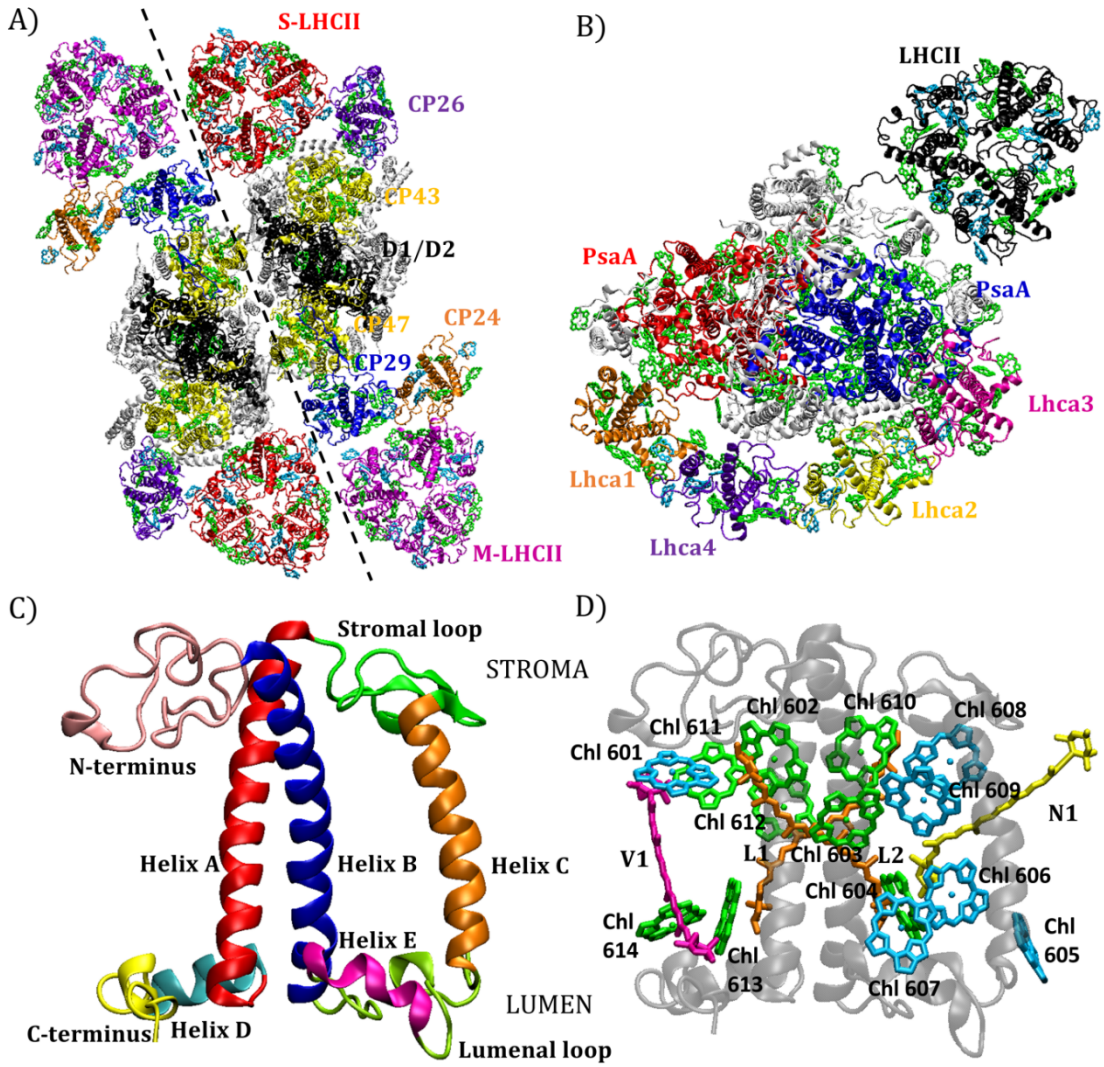


Figure 5. Architecture of LHCs and their association with the core photosystems. A) Structure of the $C_2S_2M_2$ PSII supercomplex from *Pisum sativum*[38] (PDB code: 5XNL) viewed from the stroma. Chls *a* are shown in green, Chls b in cyan. Chl-binding protein subunits are shown in different colors: black (D1/D2), yellow (CP43/CP47), blue (CP29), violet (CP26), orange (CP24), red (S-LHCII trimer), magenta (M-LHCII trimer). All other subunits are in grey. B) Structure of the PSI-LHCI-LHCII supercomplex from maize[96] (PDB code: 5ZJI) viewed from the stroma. The Chl binding subunits are shown in different colors: blue (PsaA), red (PsaB), orange (Lhca1), violet (Lhca4), yellow (Lhca2), magenta (Lhca3), and black (LHCII trimer). All other subunits are in grey. C) Side view of the LHCII monomeric protein from *Spinacia oleracea* (PDB code: 1RWT)[116], with different structural elements highlighted with different colors. D) Side view of LHCII and its pigments: the protein is shown in transparent grey, Chls *a* in green, Chls *b* in cyan, luteins in orange, neoxanthin in yellow, and violaxanthin in magenta. The Chl and Car binding sites are labeled. For clarity, only chlorin rings of the Chl pigments are shown.

Plant PSII-LHCII supercomplexes of different sizes have been purified, the largest so far being the so-called $C_2S_2M_2$ dimeric supercomplex[38,117]. In this notation, C_2 stands for the previously discussed dimeric core, while S_2 and M_2 represent LHCII trimers strongly or moderately attached to each core, respectively. $C_2S_2M_2$ also binds three monomeric antennae per core. While the overall organization of the antenna and core in $C_2S_2M_2$ are known from a longer time, near-atomic resolution structures of this supercomplex have been published only recently for both *P. sativum*[38] (Figure 5A) and *A. thaliana*[118]. All structures show that S-LHCII is located between CP26 and CP29' (the prime superscript designated a unit of the other PSII monomer) and is directly connected to the core via CP43, whereas CP29 is located between the more peripheral M-LHCII and CP47. CP24 (which is the last LHC to be evolved and is missing in algae and some gymnosperms)[119,120] is also adjacent to M-LHCII and CP29 and faces the nearby units with only Chls *b*. It is therefore predicted to be less well energetically connected to the rest of PSII. From this picture, CP29 acts as a nexus in the supercomplex, at the crossing points of a number of EET routes from the periphery to the core. Its energetics as a function of this pivotal position will be discussed in chapter 2 in more detail. The LHCII antenna markedly increases the absorption cross-section of PSII. At the same time, charge separation becomes slower because a longer time is required on average for the excitations formed in the distal antenna to reach the RCs and drive photochemistry. As a result, the average fluorescence lifetime of PSII increases from about 100 ps in dimeric cores (C_2)[91] to about 150 ps in $C_2S_2M_2$[121], and up to 300-400 ps in the thylakoid membrane[122–124], where the PSII antenna also contains extra LHCII units that are less well connected to the core. Assuming an average lifetime of 2 ns in closed state, the quantum efficiency of PSII photochemistry remains relatively high also in the membrane (dropping from 95% to about 80-85%).

In plants, the PSI monomeric core forms a supercomplex with the LHCI (LHCas) antenna complexes[37,92] (Figure 5B). Four LHCas (Lhca1-4) form a half ring around the core, each binding at a fixed position[125] and independent of light conditions[126]. The LHCas contain the most red-shifted Chls and all transfer excitations to the core, though the red-most ones (Lhca3-4) somewhat slowly. As a result, the fluorescence kinetics of PSI-LHCI supercomplexes shows two phases, a fast one (about 20 ps) due to trapping from the core and a slower one (60-100 ps) due to trapping from the red Chls of the LHCas[127]. The average lifetime is in the order of 50 ps. PSI-LHCI-LHCII supercomplexes containing an extra LHCII trimer have also been purified, with LHCII typically bound to the core on the opposite side than the LHCI antenna[96,128] (Figure 5B). The actual size of the LHCII pool connected to PSI *in vivo* is species-dependent[129] and is sensitive to the light conditions, which is used to balance the energy fluxes directed to the RCs of PSI and PSII and, therefore, their energy output (see the following section)[130]. Energy transfer from LHCII to the PSI core is also relatively fast, but the size of the LHCII pool correlates with the average PSI fluorescence lifetime[129], similar to what previously discussed for PSII supercomplexes.

LHCs are integral membrane proteins of 21-29 kDa with a highly conserved structure[37,38,93,116,131]: they consist of three transmembrane alpha helices (A-C) and two shorter amphipathic helices (D-E) exposed to the lumen (Figure 5C). The N-terminal and C-terminal domains are exposed to the stroma/lumen, respectively. The central helices represent the most rigid part of the complex and bind most pigments. Conversely, the C- and N-terminal portions, as well as the stromal and lumenal loops between helices C-A and B-C, respectively, are more flexible[132–135]. The N-terminal region has some important biological functions: in LHCII and CP29 it is the site of phosphorylation[136], and in LHCII it also contains the motif responsible for trimerization[137]. The N-terminus of CP29 is particularly long and is believed to stabilize the anchoring of CP29 to the PSII (as it extends over the stromal side of CP47 to reach the extremity of D1)[38]. It has also been recently proposed to promote interactions between stacked PSII units by extending into the stroma[138].

LHCs bind up to 15 Chls, most Chl binding sites being conserved, with possibly different Chl *a*/*b* occupancies in different LHC variants (Figure 5D). The Chls are oriented roughly perpendicular to the membrane plane and can be divided into two layers, one closer to the stroma and one the lumen. Both layers of Chls are also parallel to the membrane plane. In LHCbs, the 8 Chls on the stromal side (601-603 and 608-612, the notation is based on Liu et al.[116]) form a sort of elliptical ring, whereas those on the lumenal side form two smaller clusters (604, 606-607 plus 605 in LHCII only and 613-614; the latter two Chls are missing in CP24). When bound to the PSII supercomplex, CP29 also possesses an extra Chl (616), which is coordinated by the uniquely long N-terminus and might be relevant for strengthening its energetic connectivity with CP47[38]. Most Chls are bound to the protein via coordination of their central Mg by side chains of nucleophilic residues (usually histidine, glutamate, glutamine, or asparagine). These binding sites have been first pinpointed by mutational analysis[139–143] and later confirmed by structural data[37,38,116,131]. Other Chls are coordinated via water molecules, by the protein backbone or even by the phosphate group of a phospholipid molecule (Chl 611).

Three Car binding sites (named L1, L2 and N1) are found in all LHCs (Figure 5D). A fourth binding site, named V1, is only present in LHCII trimers but is lost in all monomeric LHCs. The two central sites – L1 and L2 – form a transmembrane cross that is intimately associated with the two central helices, assisting protein folding/stability. While L1 always binds the xanthophyll lutein, the occupancy of the L2 site varies in different LHCs (a second lutein is present in LHCII and CP26, whereas violaxanthin is found in CP29, CP24 and the LHCas). The N1 site is occupied by neoxanthin in LHCII, CP29 and CP26, and by β-carotene in CP24 and in some of the LHCas[37,38]. The N1 carotenoid is highly bent and partly extends outside the complex, into the membrane. In LHCII trimers, V1 is occupied by violaxanthin or lutein[144] and can be exchanged by other xanthophyll cycle carotenoids (antheraxanthin and zeaxanthin) in high light[145,146]. This site is located at the interface between two monomers[116], its carotenoid binding is weaker compared to other sites and is lost upon monomerization[147]. Due to its distant location from the Chls, it has been shown to be involved neither in light harvesting nor in photoprotection[144,146,148].

Carotenoid binding to LHCs is most likely contributed by a combination of hydrophilic and hydrophobic interactions, but its details are not yet understood. Only one residue has been identified so far that is important for the binding of a carotenoid, i.e. the Tyr (found in all LHCbs but CP24) forming a hydrogen bond with the –OH group of neoxanthin in N1[149]. Except for V1, all Car binding sites are involved in light harvesting and the two central ones also have a key role in the scavenging of Chl triplets[150]. Due to their central location and their stronger interactions with some crucial Chl clusters (610-611-612-L1 and 602-603-609-L2), the xanthophylls in L1 and L2 represent ideal candidates for quenching of Chl singlet excitations upon light stress[151–153].

The LHCas are structurally similar to the LHCbs, most Chl and Car binding sites being maintained. Two relevant differences are the higher relative content of Chl *a* in the LHCas and the substitution of neoxanthin by β-carotene in N1. Despite the high similarity, however, these two types of LHCs possess distinct fluorescence properties[115,154,155]. While in the LHCbs, the lowest energy pigments emit around 680 nm[156,157], the LHCas possess Chls red-shifted to different degrees, which translates into a drastic red-shift of their emission (from 690 nm in Lhca1[154] to 735 nm in Lhca4[158,159] at 77 K). As for the red Chls of PSI core, this effect has been ascribed to the presence of a mixed exciton-CT state in the excited-state manifold on these complexes[73]. The red-shifted pigments have been identified as Chls *a*603-609 via mutational analysis[141,160–162]. The reason for two highly homologue LHC architectures showing such distinct spectroscopic behaviors is not yet understood. This difference underlines how subtle changes in pigment-pigment interactions can result in a wide range of excited-state properties[163], which is essential for the fine-tuning of light harvesting and its interplay with energy dissipation.

Phycobiliproteins are organized into phycobilisomes

Phycobilisomes (PBSs) are very large water-soluble multi-protein complexes (typically 5-20 MDa) and represent the main light-harvesting complex of cyanobacteria. They are also found in the chloroplasts of red algae and glaucophytes[12,164,165]. PBSs consist of brilliantly-colored water-soluble phycobiliproteins (PBPs) and associated linker proteins[166] (Figure 6A). The former ones contain covalently-attached bilins, while the latter are generally pigment-less and direct the assembly of PBPs into PBSs. PBSs can bind several hundred pigments and, besides increasing the absorption cross-section of the photosystems, they also absorb at wavelengths (commonly between 500 and 650 nm) that are complementary to those absorbed by the Chls. PBSs are located at the stromal side of the thylakoid membrane and can transfer excitations to either PSII, PSI, or both[167–169]. The energetic coupling of the PBS to PSI and PSII seems to be dependent on the light conditions[170].

PBSs have different architectures, with highly different size, shape, and protein composition. They can be divided in five different structural classes: (I). hemi-ellipsoidal PBSs; (II) bundle-shaped PBSs; (III) hemidiscoidal PBSs; (IV) block-shaped PBSs; and (V) rod-shaped PBSs. Hemi-ellipsoidal and block-shaped PBSs are typically found in red

algae and are the largest known PBS[171,172]. The structure of the hemi-ellipsoidal-shaped PBS of the red alga *Porphyridium purpureum* has been recently solved to nearly-atomic resolution by cryo-electron microscopy[172], which represents a unique case for PBS, so far. Bundle-shaped PBSs are hexagonally packed bundles of rods that are up to 70 nm long and stand perpendicular to the membrane surface[173]. They are found in *Gloeobacter* spp., the only cyanobacteria that lack thylakoid membranes[174]. Rod-shaped PBSs are the most simple PBS architecture and occur in some well-known cyanobacteria, such as *Nostoc* sp. PCC 7120[175] and *Acaryochloris marina*[176]. The rods are composed of one or more types of PBPs and a special linker protein. The latter has a hydrophobic α-helical domain that anchors the PBS to the thylakoid membrane[175]. Most cyanobacteria have rather large hemidiscoidal PBSs of variable size (width of 45–70 nm, height of 30-50 nm and ~15 nm of depth/thickness[177,178]. Hemidiscoidal PBSs consist of a core of two, three, or five cylindrical substructures and of a number of peripheral rods that radiate from the core (Figure 6B).

Based on the absorption wavelength of their bilins, PBPs can be classified into three types: (i) phycoerythrins (PEs) are sometimes found at the core-distal ends of the rods and absorb at higher energies (λ_{max} = ~ 560 nm), (ii) phycocyanins (PCs) form the core-adjacent portions of the rods and absorb at intermediate energies (λ_{max} = ~ 620 nm), while (iii) allophycocyanins (APCs) are the major components of the core and absorb at lower energies (λ_{max} = ~ 650 nm). This organization creates a downhill energy gradient in the PBS that funnels the excitations from the periphery to the core and, from the latter to the Chls of the adjacent photosystems[179]. Although PBPs have different absorption properties, they have similar crystal structures and assemble into PBSs through a common hierarchical organization[166,180,181]. The basic assembly is a heterodimeric protomer (αβ), formed by α and β subunits with similar globin folds. αβ subunits subsequently assemble into $(\alpha\beta)_3$ trimers[182] (Figure 6A), which can stack face to face to form $(\alpha\beta)_6$ hexamers with or without linker proteins in their central cavity[166,183]. $(\alpha\beta)_3$ trimers are disk-shaped and represent the basic unit of the cylindrical substructures of the PBS[178]. The α and β subunits of APC each carry a single bilin chromophore, whereas the PC α and β subunits carry one and two chromophores, respectively[166].

The PBS core cylinders typically have four trimeric disks composed of six different APC subunits: ApcA, ApcB, ApcC, ApcD, ApcE and ApcF[165]. The most abundant units are ApcA, ApcB (α- and β- type, respectively). ApcE (α-type) has a PBP domain and a linker domain that is used to anchor the PBS to the thylakoid membrane (for this reason, ApcE is also known as the core-membrane linker PBP, L_{CM}). ApcE acts as a scaffold for the assembly of the entire core substructure and has a variable mass/size (75-120 kDa) depending on the number of cylinders forming the PBS core. Two ApcE units are required to scaffold the other components in all core types[166]. ApcC is a small (~ 10 kDa) cylinder-terminating core linker protein found at the ends of the core cylinders[178,182]. ApcF (β-type) is the assembly partner of the PBP domain of ApcE and thus there are also two copies of this protein per PBS core. Finally, ApcD (α-type) is found in one of the outer-most trimers

of each basal core cylinder[184,185]. ApcD and ApcE contain the red-most bilins, absorbing at about 665-670 nm and emitting at 680 nm, and are therefore referred as "terminal emitters"[184,185]. Mutational and structural studies have shown that the two chromophores of ApcE in the core substructure are primarily responsible for energy transfer from the PBS to PSII, while ApcD primarily mediates energy transfer from the PBS to PSI[168,169,178,186,187].

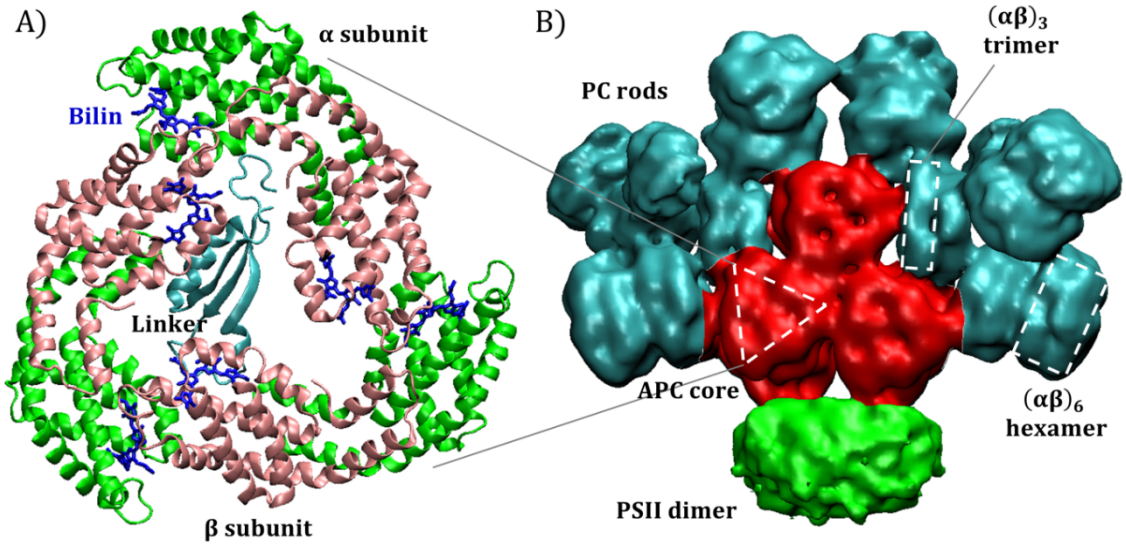


Figure 6. Phycobiliprotein and phycobilisome organizations. A) Allophycocyanin $(\alpha\beta)_3$-linker complex from *Mastigocladus laminosus*[182] (PDB code: 1B33). The type-α protein subunits are shown in green, the type-β subunits in pink, and the linker protein in cyan. Each APC subunit binds a phycocyanobilin molecule covalently (in blue). B) Electron density map of a hemidiscoidal PBS attached to a PSII dimer from *Anabaena* sp. PCC 7120[178] (Electron Microscopy Data Bank, 2822). The APC tri-cylindrical core is colored in red, the phycocyanin (PC) rods in cyan, and the PSII dimer in green. The basic structural elements are circled with white dashed lines.

Energy transfer in the PBS takes place on different timescales due to its hierarchical organization: equilibration within the same rod/cylinder is relatively fast (~ 20 ps), whereas that between core cylinders and from the peripheral rods to the core are slower (50-100 ps). Due to these processes, the average time needed to transfer excitations from the PBS to the photosystem is in the order of 100-200 ps, and the energy transfer efficiency is in the order of 90 %[165,179,188].

Flexibility of light harvesting

As mentioned in the introductory section, photosynthetic organisms have evolved a high multiplicity of antenna systems using a variety of pigments, protein units, protein environments and super-complex organizations. This versatility is needed to adapt photosynthetic light harvesting to different external conditions: by selecting the pigment

matching the available light spectrum, or by changing the antenna size (i.e. the number of pigments) to capture the optimal amount of photons, for instance[189].

The flexibility of light harvesting, however, is not limited to the differences observed between diverse species or domains. Each photosynthetic organism is subject to environmental changes throughout its lifetime, involving light quality, nutrient availability, temperature, humidity, etc. In order to face all these kinds of fluctuations, plants, algae and cyanobacteria have evolved a range of long-term and short-term responses. Most of this thesis will focus on some of these regulatory strategies, which are developed by the light-harvesting complexes or by the photosystem cores.

An additional degree of diversification is achieved by the LHCs of plants and algae. Despite showing similar spectroscopic properties, as a consequence of their common architecture, they have fine-tuned more specific features based on their function and position in the PSI or PSII antenna. This is particularly evident when considering the distinct fluorescent properties of LHCas and LHCbs (as discussed previously), but some relevant differences are also observed, for instance, between the various LHCbs. This will be the subject of chapter 2 of this thesis.

Coping with rapid changes in light intensities: non-photochemical quenching

One of the main sources of plant stress is represented by fluctuations in light intensity. These fluctuations can be relatively slow (seasonal or day/night cycles) or much faster (random dynamics in the motion of clouds, shading leaves in a canopy, etc.). Exposure to high light can saturate the electron transport chain and create a high amount of Chl excitations in the membrane that cannot be processed (quenched) by photochemistry. The resulting longer lifetime of Chl excitations would translate into a higher yield of intersystem crossing and formation of Chl triplets, which can sensitize oxygen and damage the photosynthetic machinery (protein, lipids and pigments)[190,191]. If not repaired, photodamage decreases the rate and efficiency of photosynthesis. To avoid these negative effects caused by high light exposure, plants have developed several photoprotective strategies.

Based on their timescales, we recognize: (i) long-term responses (acclimations), involving *de novo* synthesis of proteins/pigments/lipids in order to tailor the photosynthetic capacity and harvest the optimal amount of photons and (ii) short-term responses, where the unavoidable excess of absorbed energy is safely dissipated as heat[192]. This is achieved by increasing the rate (and yield) of non-radiative decay of Chl excited states at the expense of the concurrent triplet and fluorescence yields. PSII is more prone to photodamage and represent, therefore, the main target of these short-term responses referred as NPQ (non-photochemical quenching). NPQ is characterized by quenching of Chl fluorescence[193,194] and by the reduction of the excited-state lifetime of PSII[56,122,195]. The rate of NPQ is experimentally determined by comparing the fluorescence yields (or lifetimes) of closed PSII in the absence and in the presence of NPQ. Based on their origin and activation/deactivation timescales, two main NPQ components can be distinguished:

- qI indicates a reduced photochemical yield (and fluorescence quenching) of PSII caused by a photo-damaged RC or oxygen-evolving complex. Recovery from qI requires the synthesis of a new D1 unit, whose timescale (up to hours) represents the bottleneck in the relaxation kinetics of this NPQ component[190,196–198].
- qE (energy-dependent quenching) represents the fastest component in NPQ[199]: its development requires seconds to minutes[200] and (unlike qI) entirely relies on the acidification of the lumen brought by intense oxygen evolution and electron transfer in high light[193,194]. Recovery from qE follows the relaxation of the lumenal pH and is somewhat slower than qE induction, but still takes place on a timescale of minutes. The qE locus in plants has been largely ascribed to the PSII peripheral antenna[195,201–204], though a significant amount of quenching can still be observed in plants deprived of the LHCs, indicating that an additional quenching mechanism is also active in PSII cores[205,206]. A recent work has shown that trimeric LHCII is the main site of qE[206], although a role of CP29 and other minor antennae into quenching has been proposed by other studies[152,195,204,207,208].
  An essential effector of qE in higher plants is the protein PsbS[113], a member of the LHC superfamily. PsbS is a pH sensor, and its photoprotective function is triggered by the protonation of some of its lumen-exposed glutamic acids[209,210]. PsbS does not bind pigments to a significant extent[211,212], i.e. it is not a direct quencher of Chl excitations. Its precise stoichiometry and location in the PSII supercomplex, as well as its mechanism of action are currently the subject of intense study, but have been so far elusive. Another important effector of qE is the carotenoid zeaxanthin, which is obtained via a (reversible) double de-epoxidation of violaxanthin by the enzyme violaxanthin de-epoxidase[145]. This enzyme is activated by lumen acidification. The inter-conversion between zeaxanthin and violaxanthin through the mono-epoxidized form antheraxanthin is referred as xanthophyll cycle and normally requires several minutes. This is relatively slow in comparison to PsbS-mediated quenching, which is activated in a few seconds[213]. The role of zeaxanthin in qE is still debated. This carotenoid has been first suggested to be a direct quencher of Chl excitations[42,214]. Later experiments, however, have ruled out this hypothesis, as the freshly synthesized-zeaxanthin was observed to remain in the membrane rather than substitute the centrally bound violaxanthin of the (minor) LHCs[146].

Besides photo-inhibition, light fluctuations can also result in the impairment of the activities of PSI and PSII. Alterations in the energy balance between the two photosystems can also be caused by changes in the light spectrum. The reason for this is that the absorption spectra of PSI and PSII and their respective antennae do not overlap exactly, i.e. some wavelengths excite one of the two photosystems preferentially. In plants and algae, this is particularly true for far-red light ($700 < \lambda < 800$ nm; hereby referred as FRL), which is essentially absorbed by PSI only (due to the bound Chl *a* red forms). A FRL-enriched light spectrum is typically experienced by a deep-shaded leaf under a canopy (as the upper leaves absorb most of the visible light, leaving the “leftover” FRL to

the bottom leaves)[4]. To cope with relative changes in the visible/far-red light availability, plants and algae can modulate the antenna size of their PSI and PSII super-complexes, a process referred to as state transitions. State transitions are believed to take place by shifting part of the LHCII pool between the PSII-rich grana and the PSI-reach stroma lamellae and rely on LHCII phosphorylation[130].

Unlike photosynthetic organisms of green lineage, light harvesting in cyanobacteria is, to a large extent, a prerogative of the phycobilisome. As a result, NPQ and state transitions rely on completely different stimuli, effectors and mechanisms[215,216], which will not be treated in this thesis.

Structural dynamics of LHCs and its connection with fluorescence quenching

One of the most debated questions about qE concerns its detailed mechanism (both at the molecular level and in terms of the macro-organization and interaction of PSII components). The most accredited hypothesis is that the induction and relaxation of NPQ rely on conformational changes of the LHCs. Such conformational change would make the LHCs switch from a state functional for light harvesting, where Chl excitations are relatively long-lived and can be used for photochemistry, to a shorter-lived state functional for energy dissipation[151,199,217,218].

Even though the most common experimental way to structurally picture LHCs at the highest resolution is via static techniques, such as crystallography or cryo-electron microscopy, these complexes are highly dynamic structures that can access distinct conformation states[135,218–221]. This conformational flexibility is at the basis of the high tunability of their spectroscopic properties (among which the excited-state lifetime) and is supported by a variety of *in vivo*, *in vitro* and *in silico* results. LHCSR, the main pH sensor and qE effector in green algae, has been shown to accesses two distinct conformations marked by different Chl excited-state lifetimes, whose relative populations are dependent on the lumenal pH[222]. Raman signals associated with distinct Chl and Car conformations have been shown to correlate with the extent of qE both in the membrane and in clusters of LHCs *in vitro*[151,217]. T-jump IR experiments on LHCII have recently shown that the protein can re-arrange some of its structural elements following temperature variations[223]. Time-resolved fluorescence experiments on isolated LHCs in various environments reveal a multi-exponential excited-state decay, which is a witness of the multiple co-existing fluorescent (micro) states in the macroscopic sample[224–227]. Single-molecule experiments have confirmed that isolated LHCs can switch reversibly between states with different fluorescence lifetime, validating the correspondence between spectroscopic state and LHC conformation[228,229]. Even though LHCbs typically emit at 680 nm and LHCas above 700 nm, various experiments have shown that LHCbs also access red-emitting states and, vice versa, LHCas can switch to states with a fluorescence spectrum similar to an LHCb (although the occurrence of the ‘blue’ and ‘red’ states are different in the two types of complexes)[163]. Recently, molecular dynamics (MD) simulations on photosynthetic complexes have confirmed the inherent flexibility of their pigment and protein components, providing a theoretical basis for conformational

switches[135,221]. The importance of protein dynamics is also witnessed by various *in silico* studies, as calculations based on structures obtained from MD simulations have been often shown to reproduce experimental data better than calculations based on crystal/cryo-EM structures[230,231].

The above-discussed spectral heterogeneity of LHCs primarily stems from the peculiar optical properties of Chls and Cars, but what makes LHC so efficient and versatile is the embedding of these pigments in the protein matrix. If it is true that the protein primarily serves as a scaffold, arranging the chromophores in order to optimize energy transfer, this whole system cannot be regarded as a unique, static entity. LHCs are quite flexible structures and the protein dynamic environment has a large influence on the pigment geometries, excitation energies, relative distances and orientations. In this view, protein conformations shape the (excitation) energy landscape of the bound pigments and the nature of their mutual interactions, reflecting into specific energy transfer and dissipation pathways. Spectroscopy offers a unique view on these pathways, and is therefore the central investigation strategy throughout the entire thesis.

Who is the quencher in LHCs?

The energy landscape of LHCs is extremely complex due to this intricate network of pigments, energy states and interactions involved, which pose a number of experimental (and theoretical) challenges (see below). Hence, it is not so surprising that some fundamental aspects of the molecular mechanisms of light harvesting and photoprotection remain elusive. This ambiguity also concerns the quenching mechanism of LHCs. Quenching is caused by excitation energy transfer from the Chls to a short-lived non-fluorescent (dark) state that promptly dissipates the excitation energy as heat. The exact nature of the quencher, however, is still under debate and three main candidates have been proposed so far:

1) A mixed Chl-Chl exciton/CT state characterized by a shorter lifetime and/or a low oscillator strength. This mechanism has been proposed from studies on LHCII oligomers *in vitro*[72,232], as well as measurements on leaves[195]. A similar Chl-Chl exciton/CT state has also been proposed to account for the red-shifted emission of the LHCas. The latter complexes, however, have fluorescence lifetimes and oscillator strengths more typical of a light-harvesting state[127,233,234].
2) A Chl-Car CT state following Car-to-Chl electron transfer. This mechanism has been first reported for isolated LHCs containing zeaxanthin[152,214], but its relevance in physiological conditions remains debated[146,235]. More recently, Car-to-Chl electron transfer has been assigned as a quenching mechanism in LHCSR involving lutein[236], while another study has provided a general theoretical basis for it[237].
3) A Car singlet excited state. Chl-to-Car EET has been observed for the first time in LHCII aggregates[151], though those results were questioned later on due to the presence of experimental artefacts[232,235]. In the past years, this mechanism has been more convincingly assigned to the quenching of a LHC ancestor protein[238] and

non-native LHCII containing astaxanthin as the sole xanthophyll[19]. A similar process has also been proposed to account for a newly discovered sub-picosecond quenching pathway reported for LHCII embedded in lipid nanodisks[227]. The mechanism has also been addressed and supported by a number of theoretical works[239,240]. The exact nature of the Car excited state involved (whether S_1, or S*, or a different state) is still controversial. Conclusive proof in favor of this model will be given in chapter 3 of this thesis by studying the quenching mechanism of the native antenna CP29.

A variant of this mechanism is represented by the hypothesis that a Chl excited state undergoes a significant reduction in lifetime upon excitonically mixing with a Car singlet excited state[241]. According to this model, a Car-related signal should be detected immediately after Chl excitation, as opposed to the previous model, where the Car excited state should be gradually populated by EET from the Chls.

Due to their heterogeneity and the large number of pigments embedded, photosynthetic membranes are extremely challenging samples and, so far, the investigation of the molecular mechanism of qE has mostly addressed *in vitro* systems. *In vitro* conditions are indeed easier to control and allow to work on more homogeneous samples and with a lower amount of energetically connected pigments. This is why a variety of advanced spectroscopic studies has been performed so far on purified LHCs. Isolated LHCs, however, are largely found in their long-lived light-harvesting state[224,226]. For such a reason, quenching has been prevalently studied in different *in vitro* conditions such as in LHC crystals[217] and, especially, LHC oligomers, where the fluorescence lifetime is significantly shorter[242,243]. These systems have also been reported to share other spectroscopic features with qE *in vivo*[72,151,244,245]. Although these artificial samples have helped to expand our knowledge about quenching, little is known whether this knowledge can be confidently transferred to NPQ in the whole leaf. One important step towards this connection comes from achieving a unifying view of quenching in different *in vitro* systems and is provided by chapter 4 of this thesis.

Far-red light induced photoacclimation

Light-harvesting regulation by photosynthetic organisms also includes long-term responses to specific light colors. These so-called chromatic acclimations are typical of cyanobacteria, some of which have adapted to survive extreme environments, ranging from hot springs to soil crusts and even polar ice caps. The first documented chromatic acclimation in cyanobacteria was discovered over 100 years ago, when Gaidukov observed that cells from the strain *Oscillatoria sancta* were blue-green when grown in red light but red-brownish when transferred to green light[246]. This process was correctly interpreted as a modification in the pigment content to absorb a light color complementary to that of the cells. Later investigation revealed that this adaptation involves the expression of different types of phycobiliproteins (PBPs) in the two growth conditions[247].

A more recently discovered chromatic acclimation is adopted by a small group of cyanobacteria growing in particularly shaded niches such as tropical soil ecosystems, soil crusts, microbial mats, dense algal blooms, beach rocks, and stromatolites. All these environments are highly enriched in FRL. While under white light (WL) (or any color of visible light), these cells possess a photosynthetic apparatus that is highly similar to that of most cyanobacteria, when transferred to FRL they undergo a massive remodeling of the photosynthetic units referred as Far-Red Light induced Photoacclimation (in short, FaRLiP)[2]. FaRLiP is triggered by a FRL-sensing phytochrome and involves the expression of a 21-gene cluster encoding paralogs of most of the core subunits of PSI (PsaA, PsaB, PsaF, PsaI, PsaJ, and PsaL), PSII (PsbA, PsbB, PsbC, PsbD, and PsbH) and PBS (Figure 7A). From now on, the supercomplexes assembled in FRL will be referred to with the prefix FRL-, while those found in WL (or any visible light) with the prefix WL-. The FaRLiP cluster also encodes the enzyme Chl *f*-synthase, which produces Chl *f* from the oxidation of Chl *a* and is a light-activated highly-divergent paralogue of the PSII RC subunit D1[248]. It has been suggested that Chl *f* is synthesized by either a dimer of Chl *f*-synthase units[248] or by a super-rogue of the PSII core complex (D1-D2-CP43-CP47), where Chl *f*-synthase substitutes D1[249]. Another Chl variant, Chl *d*, is also synthesized upon FaRLiP, but its synthase has not been identified yet. A recent study, however, has found that the expression of all FRL-allophycocyanins is a prerequisite for the accumulation of Chl *d* in FRL-acclimated cells[250]. This and other results suggest that the cysteine-rich FRL-allophycocyanins are involved in the synthesis of Chl *d* from Chl *a*.

Both Chl *f* and Chl *d* have a red-shifted Q_y transition with respect to Chl *a* and provide FRL-cells with an increased absorption above 700 nm[23,24] (Figure 3F). WL-cells instead only contain Chl *a*. The newly synthesized Chl *f* and Chl *d* pigments represent only a minority of the total Chls in FRL-cells (~10% and 1%, respectively)[251–253], while their most-abundant Chl remains Chl *a*. Chl *f* is incorporated by both FRL-PSI (about 7-8 Chls *f* per monomer; Figure 7C)[253–257] and FRL-PSII (about 4 Chls *f* per monomer), while Chl *d* is bound to FRL-PSII only (about 1 per monomer)[253,258].

FaRLiP also involves the expression of paralogs of certain APC subunits (ApcB2, ApcD2, ApcD3, ApcD5, and ApcE2)[259,260]. However, while APCs normally absorb red light, FRL-APCs can absorb FRL photons well above 700 nm. This peculiar red-shift was initially ascribed to the presence of bilins that are not covalently bound[259,261], an occurrence that would extend their conjugation length. A more recent study, however, demonstrated that a substantial amount of far-red absorption in FRL-APCs is still carried out by covalently attached bilins, assigning their red-shift to a more planar configuration of the open-chain tetrapyrrole[262]. Together with the ApcC and ApcF that are also expressed in WL, FRL-APCs tend to form bi-cylindrical cores (FRL-BCs)[2,259,260]. In some FaRLiP strains, FRL-BCs are the only observed type of PBP association in FRL[259], while another strain has been reported to assemble FRL-BCs to peripheral rods containing PC and PE to form larger FRL-PBS[2,262]. Some organisms have also been reported to maintain WL-PBS and/or WL-photosystems for a long time after acclimating to FRL and expressing FRL-

units[258,260,263]. This might provide FaRLiP organisms with the capability of capturing visible photons by the more efficient WL- units in moments of sudden availability.

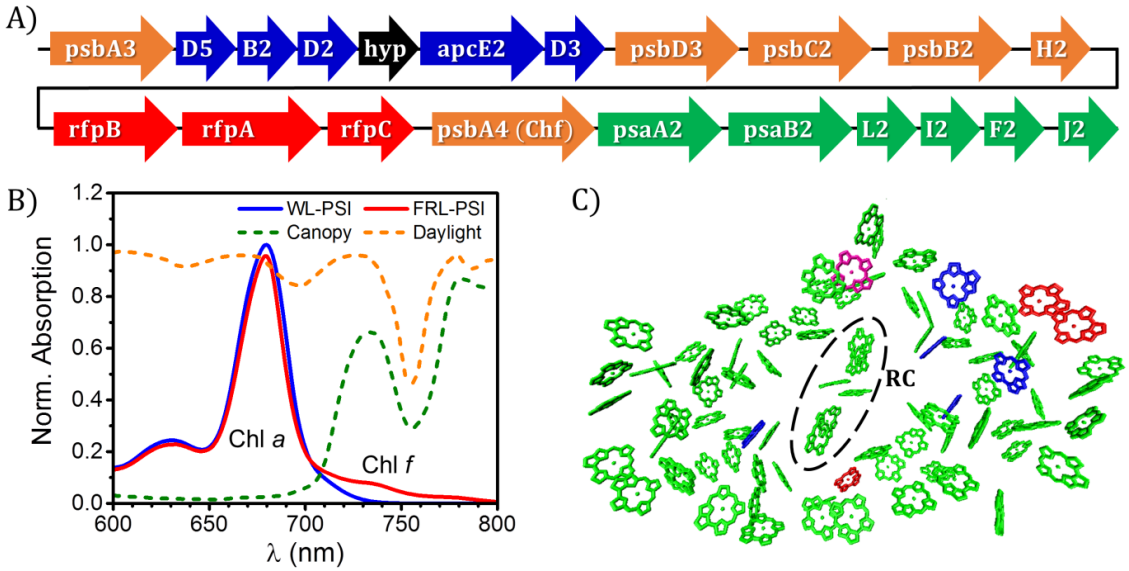


Figure 7. Remodeling of the photosynthetic apparatus upon FaRLiP. A) Physical map of the 21-gene cluster of *Leptolyngbya* sp. JSC-1 expressed in far-red light[2]. The genes encoding APC paralogs that form bicylindrical cores are shown in blue, those encoding subunits of PSI (psa) in green, and those encoding subunits of PSII (psb) in orange. The latter include the gene of Chl f synthase (Chf), a divergent paralog of PsaA (D1). The genes in red encode the knotless phytochrome photoreceptor for FRL (rfpA) and two response regulators (rfpB and rfpC). A conserved hypothetical protein (ORF) is shown in black. B) Q_y absorption spectra of WL-PSI (blue) and FRL-PSI (red) of the FaRLiP strain *Chlorogloeopsis fritschii* (normalized to the area above 600 nm; see chapter 5) and relative spectral photon irradiance of unshaded sunlight (daylight) and light filtered by a dense leaf canopy (canopy)[4]. FRL-PSI displays extra absorption above 700 nm, which makes it able to harvest light in shaded environments. C) Chl organization in the FRL-PSI from *Fischerella thermalis* by Gisriel et al.[257] (PDB code: 6PNJ). The Chl *a* molecules are shown in green, the Chl *f* binding sites of FRL-PSI according to Gisriel et al.[257] are shown in red, those according to Kato et al. (based on the structure of another FaRLiP strain, *Halomicronema hongdechloris*)[256] in blue. The only Chl *f* site common to both structures is shown in magenta. For clarity, only chlorin rings of the Chl pigments are shown.

The inserted Chls *f* extend the absorption cross-section of FRL-PSII up to 750 nm[253,258] and that of FRL-PSI up to about 800 nm[253–255] (Figure 7B). These remarkably long photon wavelengths are still able to power photochemistry in the two photosystems[253,255], implying that the PAR of FaRLiP organisms is extended well above the characteristic red limit of oxygenic photosynthesis (typically 680 nm for PSII and 700 nm for PSI). This finding has raised new questions about the energetic requirements of photosynthesis, and part of this thesis will be devoted to exploring these frontiers.

Notably, FaRLiP is not the only known strategy developed by photosynthetic organisms to make efficient use of FRL. We already discussed the role of Chl *a* red-forms, which can be found in the PSI core (and possibly, its peripheral antenna) of almost every photosynthetic organism. The red-most Chls *a* discovered so far absorb up to 740 nm at

room temperature and 760 nm at low temperatures[107]. Additionally, some algal species also possess red-shifted antennae that are functionally connected to PSII[264–267]. In all these cases, however, the red-shifted Chls only have an antenna function, whereas the RC energetics are unchanged and follow the above-mentioned red limit: the long-wavelength excitations in the antenna need energy from the environment to be transferred to the RC and drive photochemistry. Until the discovery of FaRLiP, the only known photosynthetic organism to possess red-shifted RCs, therefore breaking the red limit, was the aquatic cyanobacterium *Acaryochloris marina* (AM). AM, which also populates FRL-enriched environments, permanently uses Chl *d* as its major photosynthetic pigment, while Chl *a* is present in minor amounts (< 10%)[23,268]. The primary electron donor in AM-PSII is a Chl *d* absorbing at about 725 nm, while that of AM-PSI around 740 nm[269,270].

Based on the evidence from AM, a question arises whether Chl *f* also plays a role in photochemistry. However, contrary to AM-photosystems, where nearly all pigments absorb above 700 nm, the FRL-photosystems of FaRLiP organisms only bind a relatively small number of red-shifted pigments and locating them is far less straight forward. Consequently, while there is no doubt that some Chl *f* is present in the antenna, its location in the RC (implying an active role in charge separation) is still debated. Based on the evidence that photochemistry can still be driven by far-red photons at cryogenic temperatures, when uphill energy transferred is blocked, and on the presence of unusually red-shifted signatures in the absorption spectra of the charge-separated states, a spectroscopic study has located at least one red-shifted pigment in the RC of both FRL-PSI and FRL-PSII[253]. However, more recent structural data have excluded the presence of Chl *f* in the RC of FRL-PSI from two different FaRLiP strains[256,257] (Figure 7C). In any case, the presence of few red-shifted Chls in the antenna and (possibly) in the RC of FRL-photosystems is expected to produce substantial alterations in the energy transfer and trapping dynamics. First of all, the presence of few red-shifted Chls *f* surrounded by a majority of Chls *a* might reduce the energetic connectivity between the pigments in the antenna, and between the antenna and the RC. The red-shifted Chls in the antenna might also represent strong energy sinks competing with the RC pigments for the trapping of the excitations. Finally, a red-shifted RC might imply a reduced driving force for the primary electron transfer steps, which would decrease the rate of charge separation and increase its reversibility. All these potential pitfalls, which would translate into lower photosynthetic yields, will be investigated in chapters 5 and 6 of this thesis. The effectiveness of the energetic coupling of FRL-BCs to the FRL-photosystems will be addressed in chapter 7.

Thesis outline

As anticipated in the introduction, this thesis can be conceptually divided in two distinct sections. The first, consisting of chapters 2, 3, and 4, concerns the LHCs of plants. All the data presented in these chapters refer to a specific minor antenna, CP29. However, due to the high structural homology between different antenna complexes, the main findings of these chapters can be generalized to other members of the LHC family.

Chapter 2 targets the energy landscape of the Chls of CP29, which is investigated via spectroscopy, mutational analysis and structure-based exciton theory. Besides highlighting the main design principles of natural light harvesting, this chapter also addresses the subtle spectroscopic differences existing between different members of the LHC family, which are at the basis of their specific location and function in the photosystems. In chapter 3, using a combination of ultra-fast transient absorption, target kinetic modeling and site-directed mutagenesis, we identify the mechanism and site of energy dissipation in CP29. This result is achieved by overturning one of the paradigms of NPQ, that is, the investigation of quenching *in vitro* requires enhancing the photoprotective state of the LHCs by aggregation or crystallization. The mechanism responsible for energy dissipation was indeed studied in detergent-solubilized monomeric CP29 by taking advantage of a subpopulation of quenched complexes that are in equilibrium with the majority of unquenched units. The findings of chapter 3 leave two open questions: is the photoprotective mechanism identified in solubilized CP29 monomers also functional in other conditions (and, most importantly, in the membrane)? And if so, how are the light-harvesting and quenched states regulated when the environment changes? These two points are addressed in chapter 4, where we investigate the quenching mechanism active in CP29 aggregates *in vitro* via time-resolved fluorescence and advanced kinetic modeling. The oligomeric state of the complex is also of relevance because it mimics the densely packed clusters of photosynthetic units typically observed in the thylakoid membrane.

The second part of the thesis, including chapters 5, 6, and 7, focuses on the FRL-photosynthetic units of FaRLiP cyanobacteria. The major experimental technique adopted in this section is time-resolved fluorescence and all studies are performed on multiple FaRLiP strains to verify the variability of their acclimation responses. In chapter 5 we characterize the Chl *f*-containing photosystems both *in vitro* and *in vivo* and assess the effect of Chl *f* insertion on the timescales of energy transfer and charge separation. One of the outcomes of this chapter is that, while Chl *f* insertion reduces the efficiency of FRL-PSII, the photochemical yield of FRL-PSI remains almost as high as that of WL-PSI. This result is somewhat astonishing when considering that FRL-PSI can harvest and process excitations up to 800 nm, 100 nm above the red limit (corresponding to an uphill energy gap of about 200 meV, i.e. about 8 times the thermal energy at room temperature). Chapter 6 is therefore devoted to a deep understanding of the energy transfer and trapping mechanisms that make FRL-PSI so efficient despite its apparently unfavorable energetics. The last chapter of the thesis focuses on the FRL-BCs of FaRLiP cyanobacteria and their connectivity to FRL-photosystems. Throughout these chapters, the spectroscopic data are often combined with the available structural data to rationalize our findings.